Morris Kline

El pensamiento matemático de la antigüedad a nuestros días



Vols 1, 2 y 3

Edición impresa:
Alianza Editorial
Primera edición 1972

TALLERES ESTUDIANTILES
CIENCIAS
UNAM

Edición digital:



El pensamiento matemático
de la Antigüedad
a nuestros días, I

Alianza Universidad

Morris Kline

El pensamiento matemático de la Antigüedad a nuestros días, I

Versión española de
Mariano Martínez, Juan Tarrés, Alfonso Casal

Coordinación y revisión de Jesús Hernández

Alianza
Editorial

Título original: Esta obra ha sido publicada por Oxford University Press, Inc.
bajo el título *Matematical thought from ancient to modern times*

Calle Milán, 38, 28043 Madrid; teléf. 300 00 45
I.S.B.N.: 84-206-2957-X (O.C.)
I.S.B.N.: 84-206-2715-1 (Tomo I)
Depósito legal: M. 26.407-1992
Fotocomposición: EFCA, S. A.
Dr. Federico Rubio y Galí, 16. 28039 Madrid
Impreso en Lavel. Los Llanos, nave 6. Humanes (Madrid)
Printed in Spain

INDICE



A mi esposa, Helen Mann Kline

PROLOGO

> Si queremos prever el futuro de la matemática, el camino adecuado para conseguirlo es el de estudiar la historia y el estado actual de esta ciencia.
>
> Henri Poincaré

Este libro trata de los descubrimientos y desarrollos matemáticos más importantes llevados a cabo desde la Antigüedad hasta las primeras décadas del siglo XX. El objetivo perseguido es el de presentar las ideas centrales, poniendo un énfasis especial en aquellas corrientes de desarrollo que se han mostrado como las más importantes a lo largo de los principales períodos de la historia de la matemática, y que han ejercido una influencia destacada orientando y dándole forma a la actividad matemática posterior. También se ha prestado una gran atención al concepto mismo de matemática, siguiendo los cambios que ha experimentado este concepto a lo largo de los diferentes períodos, así como a la idea que han ido teniendo los matemáticos de su propia actividad.

Este libro debe ser considerado simplemente como un panorama general de la historia de la matemática. Si uno se para a pensar que las obras de Euler superan los setenta volúmenes, las de Cauchy tienen veintiséis y las de Gauss doce, fácilmente puede caer en la cuenta de que una obra como ésta, en un solo volumen, no puede tener pretensiones de presentar una exposición completa. En algunos capítulos de este libro presentamos solamente unas pocas muestras de lo que se creó en los campos correspondientes, aunque confiamos en que

estas muestras sean las más representativas. Por otra parte, al citar teoremas u otros resultados hemos omitido a menudo condiciones menores que se necesitarían para ser estrictamente correctos, con el fin de centrar la atención en las ideas principales. Por restringida que pueda parecer esta obra, creemos haber conseguido presentar una cierta perspectiva de la historia completa de la matemática.

El libro está organizado subrayando más bien los temas matemáticos importantes que los hombres que los desarrollaron. Cierto es que toda rama de la matemática lleva el sello de sus fundadores, y que los grandes hombres han jugado papeles decisivos al determinar el curso a seguir por la matemática, pero son sus ideas lo que queremos presentar; las biografías se considerarán como totalmente subordinadas. Hemos seguido, a este respecto, el consejo de Pascal: «Cuando citemos autores, citaremos sus demostraciones, no sus nombres.»

Por razones de coherencia, en especial para el período posterior al 1700, hemos tratado cada desarrollo en el momento en que alcanza su madurez, se destaca y ejerce su influencia sobre otros campos de la matemática. Así, por ejemplo, la geometría no euclídea aparece expuesta en el siglo XIX, a pesar de que la historia de los esfuerzos por demostrar o sustituir el axioma euclídeo del paralelismo se remonta a la época inmediatamente posterior a Euclides. Naturalmente, ha habido muchos temas que han aparecido recurrentemente en distintos períodos.

Con objeto de mantener el material dentro de límites razonables, hemos tenido que ignorar varias civilizaciones como la china [1], la japonesa y la maya, dado que su obra prácticamente no tuvo impacto sobre las corrientes centrales del pensamiento matemático. Por otra parte, a algunas teorías matemáticas como la teoría de probabilidades y el cálculo de diferencias finitas, que tienen hoy una gran importancia pero que no jugaron un papel tan importante en el período que aquí consideramos, se les ha dedicado poca atención. El enorme desarrollo de las últimas décadas nos ha obligado a incluir únicamente las creaciones del siglo XX que adquirieron su importancia dentro del período mencionado. Seguir a lo largo del siglo XX el desarrollo de temas tales como la teoría de ecuaciones diferenciales ordinarias o el

[1] Puede verse una buena exposición de la historia de la matemática china en el libro de Joseph Needham *Science and Civilization in China*, Cambridge Univ. Press, 1959, vol. 3, págs. 1-168.

cálculo de variaciones exigiría echar mano de materiales muy especializados que sólo tienen interés para los investigadores en esos campos concretos y alargaría excesivamente el libro. Aparte de estas últimas consideraciones, hay que añadir que resulta muy difícil evaluar objetivamente y sobre la marcha la importancia de muchos de los desarrollos más recientes. Precisamente la historia de la matemática nos enseña que muchos temas que provocaron un enorme entusiasmo y atrajeron la atención de los mejores matemáticos terminaron cayendo en el olvido. No hay más que recordar la afirmación de Cayley en el sentido de que la geometría proyectiva es toda la geometría, o la de Sylvester de que la teoría de invariantes algebraicos resume todo lo valioso de la matemática. En realidad, una de las preguntas más interesantes a las que viene a responder la historia es la de qué es lo que logra sobrevivir en la matemática; la historia hace, ciertamente, su propia y fundada evaluación.

De los lectores que tengan incluso unos conocimientos básicos de las docenas de campos más importantes no se puede esperar que conozcan lo esencial de todos estos desarrollos. Por tanto, y excepto en algunos temas muy elementales, se explica también el contenido de aquellos cuya historia se está estudiando, unificando así en cierto modo la exposición con la historia. Estas explicaciones de las diversas teorías pueden no llegar a clarificarlas completamente, pero deberían dar al menos una idea de su naturaleza. Consecuentemente, este libro puede servir en cierto sentido como una introducción histórica a la matemática; este enfoque constituye ciertamente uno de los mejores procedimientos para llegar a entender y apreciar correctamente una teoría.

Esperamos que este libro sea útil tanto para matemáticos profesionales como en formación. El profesional se ve hoy obligado a dedicar tanto tiempo y energías a su especialidad, que tiene pocas oportunidades de familiarizarse con su historia. Y, sin embargo, este marco histórico es muy importante. Las raíces del presente se hunden profundamente en el pasado, y casi nada de ese pasado resulta irrelevante para el hombre que trata de entender cómo el presente llegó a ser lo que es. Por otra parte, la matemática, pese a su proliferación en cientos de ramas, tiene su unidad propia y sus metas y problemas importantes y, salvo que los diversos campos contribuyan decididamente al núcleo de la matemática, corren el peligro de volverse estériles. La manera más segura de combatir los peligros que amenazan nuestra fragmentada ciencia quizás sea la de llegar a cono-

cer los logros, tradiciones y objetivos de la matemática en el pasado, para poder dirigir las investigaciones por vías fructíferas. Como muy bien dijo Hilbert: «La matemática es un organismo para cuya fuerza vital es condición necesaria la unión indisoluble de sus partes.»

Para los estudiantes de matemáticas este libro puede presentar otro tipo de interés. Los cursos usuales presentan teorías matemáticas que parecen tener poca relación unas con otras. La historia puede dar la perspectiva global del tema y relacionar las materias de los cursos no sólo unas con otras sino también con las líneas centrales del pensamiento matemático.

Asimismo, dichos cursos también resultan engañosos por otro motivo básico: en ellos se da una presentación de una teoría organizada lógicamente, que deja la impresión de que los matemáticos han avanzado de un teorema al siguiente de una manera casi natural, que pueden superar cualquier dificultad, y que las teorías están ya completamente trilladas y acabadas. La imponente sucesión de teoremas hunde en la miseria al alumno, especialmente si está empezando a estudiar la materia.

La historia, por el contrario, nos enseña que el desarrollo de cualquier rama de la matemática se ha llevado a cabo de una manera gradual, a base de resultados que solían provenir de diversas direcciones. También nos enseña que a menudo se han necesitado décadas, e incluso cientos de años, de esfuerzos antes de conseguir algún progreso de importancia. Y en lugar de la impresión de que las teorías están ya completamente trilladas y terminadas, uno se encuentra con que, a menudo, lo que se ha conseguido es simplemente un punto de partida, con que hay que rellenar aún muchos huecos, o con que todavía quedan por hacer las generalizaciones realmente importantes.

Las cuidadas y ordenadas exposiciones que se hacen en los cursos habituales no muestran en absoluto los conflictos del proceso creativo, las frustraciones, y el largo y arduo camino que los matemáticos han tenido que recorrer para llegar a construir una estructura importante. Siendo consciente de esto, el estudiante no sólo logrará un conocimiento mejor, sino que sacará de ahí el valor necesario para seguir atacando con tenacidad sus propios problemas, y no se desanimará por las deficiencias de su propio trabajo. Realmente, el conocimiento de cómo han avanzado los matemáticos dando traspiés, a veces en la oscuridad más absoluta, hasta llegar a reunir las piezas individuales de sus resultados, debería animar a cualquier principiante en la investigación.

Para cubrir el extenso período que pretende describir este libro, hemos tratado de seleccionar las fuentes más fiables. Para la época anterior al cálculo infinitesimal, estas fuentes, tales como el libro de T. L. Heath *A History of Greek Mathematics*, son obviamente fuentes secundarias, y en esos casos hemos utilizado varias de ellas y no sólo una. Para los desarrollos posteriores, casi siempre se ha podido ir directamente a las obras originales, que afortunadamente pueden encontrarse en las revistas o en las obras completas de los matemáticos más eminentes. También hemos visto facilitada nuestra labor por los numerosos informes y resúmenes de investigaciones que se encuentran a menudo en las ediciones de obras completas. Hemos tratado de dar referencias concretas de todos los resultados importantes, pero hacerlo así para todos los teoremas habría supuesto una confusa masa de referencias y un consumo de espacio que es mejor dedicarlo a la exposición misma.

Las fuentes utilizadas se indican en las bibliografías de los finales de capítulo; el lector interesado puede obtener en dichas fuentes mucha más información de la que hemos extractado aquí. Estas bibliografías incluyen también muchas referencias que no se podrían considerar como fuentes; sin embargo, se ha considerado interesante incluirlas bien porque ofrecen información adicional, porque el nivel de la presentación puede ser útil para algunos lectores, o porque pueden ser más fácilmente accesibles que las fuentes originales.

Quiero expresar mi gratitud a mis colegas Martin Burrow, Bruce Chandler, Martin Davis, Donald Ludwig, Wilhelm Magnus, Carlos Moreno, Harold N. Shapiro y Marvin Tretkoff, que respondieron a numerosas preguntas, leyeron muchos capítulos y ejercieron una valiosa crítica. Un agradecimiento muy especial debo a mi esposa Helen por su crítica del manuscrito, su extensa comprobación de nombres, fechas y fuentes, así como por su cuidadosa lectura de las pruebas de imprenta. De gran ayuda resultó la labor de Mrs. Eleanore M. Gross, que mecanografió todo el texto. Por último, quiero expresar también mi gratitud a la dirección y equipo de la Oxford University Press por su escrupulosa edición de este libro.

Morris Kline
Nueva York
Mayo 1972

Capítulo 1
LA MATEMATICA EN MESOPOTAMIA

La lógica puede permitirse ser paciente, puesto que es eterna.

OLIVER HEAVISIDE

1. ¿Dónde tuvo su origen la matemática?

La matemática, entendida como disciplina racional bien organizada e independiente, no existía antes de que entraran en escena los griegos de la época clásica, que va más o menos del 600 al 300 a. C. Hubo, sin embargo, algunas civilizaciones anteriores en las que se desarrollaron los orígenes o rudimentos primarios de la matemática. Muchas de las civilizaciones primitivas no llegaron más que a distinguir entre uno, dos y muchos, mientras que otras consiguieron acceder a números realmente grandes e incluso fueron capaces de operar con ellos. Otras aún llegaron a reconocer los números como conceptos abstractos, adoptando palabras especiales como nombres para cada uno de ellos y símbolos concretos para representarlos, e incluso introdujeron el uso de una base como el diez, el veinte o el cinco para representar una unidad de orden superior al ir contando. También nos encontramos con las cuatro operaciones aritméticas elementales, si bien restringidas a números no muy grandes, y con la idea de fracción, que solía limitarse, sin embargo, a 1/2, 1/3 y otras análogas, expresadas mediante palabras. Se reconocieron además las ideas geométricas más sencillas, como la recta, el círculo, los ángulos,

etcétera. No deja de ser interesante hacer notar que el concepto de ángulo probablemente surgiera de la observación de los distintos ángulos que pueden formar el muslo y la pierna de una persona, o su brazo y su antebrazo, porque en muchas lenguas se denomina a un lado de un ángulo con la palabra «brazo» o «pierna». En español, por ejemplo, se habla a veces de los brazos de un ángulo recto. Las aplicaciones de la matemática en estas civilizaciones primitivas se limitaron a cálculos comerciales muy sencillos, al cálculo aproximado de áreas de campos, a la decoración geométrica de la cerámica, al diseño de dibujos para reproducirlos repetidamente en los tejidos, y al registro y medida del tiempo.

Hasta que llegamos a la matemática de los babilonios y de los egipcios de hacia el año 3000 a. C., no encontramos ningún otro progreso matemático. Desde que los pueblos primitivos decidieron establecerse sedentariamente en una zona concreta, construyendo viviendas y dedicándose a la agricultura y a la domesticación de animales hacia el 10000 a. C., podemos ver lo lentamente que fue dando sus primeros pasos la matemática más elemental. Por otra parte, la existencia de buen número de civilizaciones sin matemáticas de las que podemos hablar nos muestra lo diseminado que estuvo antiguamente el cultivo de esta ciencia.

2. La historia política de Mesopotamia

Los babilonios fueron los primeros de estas dos antiguas civilizaciones en contribuir al desarrollo de las corrientes centrales de la matemática. Nuestros conocimientos acerca de las civilizaciones antiguas del Próximo Oriente, y de Babilonia en particular, son en su mayor parte el resultado de la investigación arqueológica de los últimos cien años, y por este motivo dichos conocimientos son bastante incompletos y sujetos a correcciones y modificaciones según se vayan haciendo nuevos descubrimientos. El adjetivo «babilónico» se aplica, abusando un tanto del lenguaje, a toda una serie de pueblos que ocuparon, simultáneamente o de manera sucesiva, la región comprendida entre los ríos Eufrates y Tigris y sus alrededores, región conocida como Mesopotamia y que hoy forma parte del Estado moderno de Irak. Estos pueblos vivieron en una serie de ciudades, a veces incluso políticamente independientes unas de otras, tales como Babilonia, Ur, Nippur, Susa, Aššur, Uruk, Lagash, Kish y otras.

Hacia el 4000 a. C. se instalaron en el sudeste de Mesopotamia los sumerios, distintos étnicamente de los semitas y de los pueblos indogermánicos posteriores. Su capital fue Ur, y el territorio que ocuparon se llamó Sumer. Aunque la cultura que desarrollaron los sumerios alcanzó su apogeo hacia el 2250 a. C., antes incluso, hacia el 2500 a. C., fueron dominados políticamente por los acadios, un pueblo semita cuya capital era Accad, gobernados en esa época por el rey Sargón, y así la brillante cultura sumeria quedó fusionada con la acadia, que la asimiló. Un período de alto nivel cultural se produjo durante el reinado del rey Hammurabi (hacia el 1700 a. C.), bien conocido como autor y promulgador de un famoso código legal.

Hacia el año 1000 a. C., nuevas migraciones y la introducción del hierro trajeron consigo nuevos cambios, y más tarde, durante el siglo VIII a. C., la región fue controlada por los asirios, que procedían de la zona montañosa del alto Tigris. Por lo que sabemos, los asirios no añadieron nada nuevo a la cultura anterior, y un siglo más tarde vemos el imperio asirio compartido por los caldeos y los medos, estos últimos muy próximos étnicamente a los persas, que vivían más al Este. A este período de la historia de Mesopotamia (siglo VII a. C.) se le suele llamar período caldeo. El Próximo Oriente cayó en poder de los persas, con el rey Ciro, hacia el 540 a. C., y algunos matemáticos persas de la época, tales como Nabu-Rimanni (ca. 490 a. C.) y Kidinu (ca. 480 a. C.) llegaron a ser conocidos por los griegos.

El año 330 a. C., Alejandro Magno, el gran general griego, conquistó Mesopotamia, y al período que va del 300 a. C. a los comienzos de nuestra era se le suele llamar período seléucida, del nombre del general griego que fue el primero en controlar la región a la muerte de Alejandro en el verano del año 323 a. C. Sin embargo, para entonces ya se había producido el florecimiento de la matemática griega, y desde la época de Alejandro hasta mediados del siglo VII d. C., en que entraron en escena los árabes, fue la influencia griega la que predominó en el Próximo Oriente. En cualquier caso, la mayor parte de las contribuciones de los babilonios a la matemática son muy anteriores al período seléucida.

A pesar de los numerosos y frecuentes cambios de gobernantes en Mesopotamia, en el desarrollo de la matemática se dio una continuidad notable de conocimientos, tradición y práctica desde los tiempos más antiguos hasta la época de Alejandro por lo menos.

3. Los símbolos numéricos

La principal fuente de información que tenemos sobre la civilización y la matemática babilónica, tanto de la antigua como de la más reciente, la constituyen los textos grabados en tablillas de arcilla. Estos textos se escribían sobre las tablillas cuando la arcilla aún estaba blanda, y a continuación se cocían en hornos o simplemente se endurecían al sol. Este procedimiento ha garantizado la buena conservación de las que no han resultado destruidas. Estas tablillas datan principalmente de dos períodos: algunas de hacia el 2000 a. C., y en mayor cantidad del período que va desde el 600 a. C. al 300 d. C. Las del primer período son las más importantes por lo que se refiere a la historia de la matemática.

La lengua y la escritura utilizadas en las tablillas del período más antiguo es el acadio, que se superpuso al tipo de lenguaje y escritura sumerios, más antiguo aún, como hemos dicho. Las palabras de la lengua acadia consistían en una o más sílabas, y cada sílaba venía representada por un grupo de signos que se reducían esencialmente a pequeños segmentos rectilíneos. Los acadios utilizaban para escribir un prisma de sección triangular, que apoyaban sobre la tablilla en una posición inclinada, produciendo así unas señales en forma de «cuña» orientadas en distintas direcciones. Esta escritura recibió más tarde el nombre de «cuneiforme», de la palabra latina *cuneus*, que significa «cuña».

La aritmética alcanzó su más alto grado de desarrollo en la civilización babilónica durante el período acadio. Los números naturales se escribían de la manera siguiente:

1 2 3 4 5 6 7

8 9 10 11 12 20 30 40 50

60 70 80 120 130

Las características más sorprendentes del sistema numérico babilónico son el principio de notación posicional y la base 60.

Al principio, los babilonios no tenían ningún símbolo para indicar la ausencia de unidades de un orden o posición cualquiera y, por lo tanto, sus numerales podían resultar ambiguos; así, por ejemplo, podía significar 80 ó 3.620, según que el primer símbolo signifique 60 ó $60^2 = 3.600$. A menudo se utilizó un espacio vacío más extenso que lo normal para indicar la ausencia de unidades de una posición dada, pero evidentemente esto podía ser mal interpretado y resultar confuso. Durante el período seléucida se introdujo un símbolo especial de separación para indicar una posición vacía. Así, la expresión representaba el número $1 \cdot 60^2 + 0 \cdot 60 + 4 = 3.604$. Sin embargo, incluso en este período no se utilizó ningún símbolo para indicar una o más posiciones vacías por el extremo derecho del número, como, en nuestra notación, 20. En ambos períodos, pues, para saber el verdadero valor de una expresión numérica había que recurrir al contexto en el que aparecía o se utilizaba, lo cual, evidentemente, podría aclarar casi cualquier duda que se presentase.

Pero los babilonios también utilizaron el principio de notación posicional para representar las fracciones, lo que constituye sin duda el aspecto más notable y útil de su invención. Así, por ejemplo, , entendido como fracción, representaría 20/60, y , como fracción, podría representar 21/60 o bien $20/60 + 1/60^2$. La ambigüedad mencionada más arriba en cuanto a las posiciones vacías seguía presentándose aquí, lógicamente.

Algunas de las fracciones más sencillas venían representadas por símbolos especiales. Así, nos encontramos con para 1/2, para 1/3, y para 2/3. Estas fracciones especiales, 1/2, 1/3, y 2/3, eran para los babilonios «totalidades», en el sentido de medidas de cantidades y no de divisiones de la unidad en partes, aunque, naturalmente, debieron surgir como medidas de cantidades que guardaban esas relaciones respectivas con otra cantidad tomada como unidad. Así, nosotros mismos podemos escribir 10 céntimos como 1/10 de peseta, pero seguir pensando en este 1/10 como una unidad en sí misma.

En realidad, los babilonios no utilizaron exclusivamente la base 60. A veces, sobre todo para representar años, escribían cosas como 2 *me* 25, donde la palabra *me* representa 100, es decir, que se trata del año 225. De la misma manera se usó *lîmu* para 1.000, generalmente en

textos no matemáticos, aunque a veces aparezca incluso en algunos textos matemáticos del período seléucida. También se pueden encontrar a veces mezclados el 10 y el 60, como en 2 *me* 1, 10, que significa $2 \cdot 100 + 1 \cdot 60 + 10 = 270$. Sistemas mixtos, utilizando una amplia variedad de unidades de diversos órdenes, tales como 60, 24, 12, 10, 6 y 2, se usaron para fechas, áreas, medidas de peso, monedas, etc., más o menos como nosotros usamos 12 para las horas, 60 para los minutos y los segundos, 10 para contar, etc. El sistema babilónico, en el fondo lo mismo que el nuestro, constituía el resultado de diversas costumbres históricas y regionales. Sin embargo, en los textos matemáticos y astronómicos utilizaron casi exclusivamente la base 60.

No sabemos con seguridad cómo llegó a generalizarse el uso de esta base 60. Una de las posibles explicaciones sugiere que pudo venir aconsejada por los diferentes sistemas de medidas de peso; supongamos que tenemos un sistema de medidas de peso con valores que están entre sí en las relaciones

$$1/2,\ 1/3,\ 2/3,\ 1,\ 10$$

y supongamos que hay otro sistema con una unidad distinta pero las mismas relaciones anteriores, y que razones de tipo político o social aconsejan fusionar los dos sistemas (como si se tratara de metros y pies, por ejemplo). Si la mayor de las dos unidades fuera 60 veces la menor, entonces 1/2, 1/3 y 2/3 de la mayor serían múltiplos enteros de la menor, y así podría haberse adoptado la unidad mayor por resultar tan conveniente.

En cuanto a los orígenes de la notación posicional, hay al menos dos explicaciones posibles. En un sistema antiguo de escritura numérica, 1 multiplicado por 60 venía representado por un símbolo 𒁹 más grande que el mismo símbolo para el 1. Ahora bien, al irse simplificando la escritura, el 𒁹 grande se fue reduciendo de tamaño, aun conservando su valor usual de 60 y, por lo tanto, acabó representando un múltiplo cualquiera de 60, según su posición dentro del numeral. Otra explicación posible viene sugerida por el sistema de monedas utilizado. Un talento y 10 mana pudo escribirse como 𒁹𒌋 donde 𒁹 significaba un talento, que equivalía a 60 mana. Nosotros seguimos haciendo lo mismo al escribir 1,20 pesetas, donde el 1 representa en realidad 100 céntimos. De esta manera, el sistema para escribir

cantidades de dinero pudo haber sido adoptado en la aritmética con toda generalidad.

4. Las operaciones aritméticas

En el sistema babilónico los símbolos para el 1 y para el 10 eran los símbolos básicos; los números del 1 al 59 se construían combinando más o menos de estos símbolos, de manera que las operaciones de sumar y restar se reducían a añadir o quitar símbolos. Para representar la suma los babilonios reunían las dos expresiones en una sola, como en , que significa 10 + 6 = 16. La resta se solía indicar por el símbolo ; así, representa 40 − 3. En los textos astronómicos tardíos aparece a veces la palabra *tab* para designar la operación de sumar.

También efectuaban los babilonios multiplicaciones de números enteros: multiplicar por 37, por ejemplo, suponía multiplicar por 30, luego por 7, y sumar los resultados. El símbolo específico para la multiplicación era el , que se pronunciaba *a-ra* que significa «ir».

Para dividir un número entero por otro los babilonios procedían de la manera usual, y dado que dividir por un entero a es lo mismo que multiplicar por su inverso $1/a$, en este punto se presentaban inevitablemente las fracciones. Los babilonios representaban los inversos como «decimales» sexagesimales y, salvo en los pocos casos que hemos mencionado más arriba, no utilizaban símbolos especiales para las fracciones. Para ello se habían construido tablas que mostraban cómo expresar números del tipo $1/a$ en forma sexagesimal finita, donde $a = 2^{\alpha}3^{\beta}5^{\gamma}$. Sólo en algunas tablillas se dan valores aproximados para 1/7, 1/11, 1/13, etc., porque estas fracciones conducen a expresiones sexagesimales infinitas que se repiten periódicamente. Cuando aparecían en los problemas más antiguos fracciones con denominadores que incluían algún factor primo distinto de 2, 3 ó 5, entonces el mismo factor molesto aparecía también en el numerador y simplemente se cancelaba uno con otro.

Los babilonios utilizaron sistemáticamente estas tablas de inversos. Dichas tablas nos muestran textos como el siguiente, por ejemplo:

igi 2 *gál-bi* 30	*igi* 8 *gál-bi* 7, 30
igi 3 *gál-bi* 20	*igi* 9 *gál-bi* 6, 40
igi 4 *gál-bi* 15	....
igi 6 *gál-bi* 10	*igi* 27 *gál-bi* 2, 13, 20

que significa, obviamente, que 1/2 = 30/60, 1/3 = 20/60, etc. El significado exacto de las palabras *igi* y *gál-bi* nos es desconocido. Las fracciones sexagesimales, es decir, los números menores que 1 expresados en términos de las sucesivas potencias de 60, 60^{-1}, 60^{-2}, etc., pero en las que los denominadores simplemente se sobreentendían, se siguieron usando por los griegos, como Hiparco y Ptolomeo, así como en la Europa renacentista hasta finales del siglo XVI, en que se vieron desplazados al fin por los «decimales» en base 10.

Los babilonios disponían también de tablas de cuadrados, raíces cuadradas, cubos y raíces cúbicas. Cuando la raíz en cuestión era un número entero se daba su valor exacto, y para las demás el valor sexagesimal correspondiente era sólo aproximado, desde luego, puesto que los números irracionales no se pueden expresar con un número finito de cifras decimales ni sexagesimales. Sin embargo, no hay ninguna evidencia en absoluto de que los babilonios fueran conscientes de este importantísimo hecho, sino que lo más plausible es que creyeran que los irracionales también se podían expresar de manera exacta en forma sexagesimal, prolongando la expresión hasta donde fuera necesario. Una excelente aproximación babilónica a $\sqrt{2}$ da como valor $\sqrt{2} = 1{,}414213...$ en vez del correcto 1,414214...

Las raíces aparecen, por ejemplo, al calcular la diagonal d de un rectángulo de altura h y base w. En un problema se pide calcular la diagonal de una puerta rectangular de altura y anchura dadas; la respuesta viene dada sin más explicaciones, y se reduce a utilizar la fórmula aproximada para la diagonal d,

$$d \approx h + \frac{w^2}{2h}$$

Esta fórmula da una buena aproximación de d si $h > w$. Así, para el caso $h > w$, como ocurre en un problema, se puede ver que el resultado es bastante razonable, observando que

$$d = \sqrt{h^2 + w^2} = h\sqrt{1 + \frac{w^2}{h^2}} = h\left(1 + \frac{w^2}{h^2}\right)^{1/2}.$$

Si desarrollamos la última expresión aplicando el teorema binomial y nos quedamos con los dos primeros términos solamente, obtenemos exactamente la fórmula anterior. Otros resultados aproximados para problemas de raíces cuadradas provienen seguramente de tablas numéricas, tan frecuentes en la matemática babilónica.

5. El álgebra babilónica

Aparte de las tablas, que nos suministran abundante información sobre el sistema numérico y las operaciones aritméticas babilónicas, hay otras con textos que contienen problemas algebraicos y geométricos. Un problema típico del álgebra babilónica más antigua pide hallar un número tal que sumado a su inverso dé un número dado. En notación moderna podemos escribir que lo que buscaban los babilonios eran dos números x y $\bar{x}$ tales que

$$\begin{cases} x\bar{x} = 1 \\ x + \bar{x} = b \end{cases}$$

Estas dos ecuaciones dan como resultante una ecuación cuadrática en x,

$$x^2 - bx + 1 = 0.$$

Los babilonios calculaban $b/2$, luego $(b/2)^2$ y por último $\sqrt{(b/2)^2 - 1}$; entonces

$$\frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - 1} \quad \text{y} \quad \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - 1}$$

son los valores buscados de x y $\bar{x}$. Los babilonios disponían, en efecto, de la fórmula para resolver ecuaciones cuadráticas. Otros problemas, como el de hallar dos números, dados su suma y su producto, se reducían al caso anterior. Dado que los babilonios no conocían los números negativos, nunca se consideran las posibles raíces negativas de las ecuaciones de segundo grado. A pesar de que en las tablillas sólo aparecen ejemplos concretos, la mayoría de ellos sin duda intentaba ilustrar un método general para las ecuaciones cuadráticas;

los casos de problemas algebraicos más complicados se reducían por medio de transformaciones a otros más sencillos.

Los babilonios llegaron a resolver problemas concretos que conducían a sistemas de cinco ecuaciones con cinco incógnitas, e incluso hay un problema, que aparece en el contexto de una corrección de observaciones astronómicas, que conduce a un sistema de diez ecuaciones con diez incógnitas, la mayor parte de ellas lineales. La solución del sistema utiliza un método especial de ir combinando las ecuaciones hasta llegar a calcular los valores de las incógnitas.

Los problemas algebraicos aparecen formulados y resueltos de una manera completamente verbal, sin utilizar símbolos especiales. A menudo aparecen las palabras *us* (longitud), *sag* (anchura) y *aša* (área) utilizadas para representar las incógnitas, no porque dichas incógnitas representen necesariamente tales cantidades geométricas, sino probablemente porque muchos problemas algebraicos surgieron de situaciones geométricas y la terminología geométrica acabó por imponerse como terminología corriente. Un ejemplo de la manera en que se utilizaban estos términos para representar las incógnitas, así como de la forma en que aparecen formulados los problemas, puede ser el siguiente: «He multiplicado la longitud por la anchura y el área es 10. He multiplicado la longitud por ella misma y he obtenido un área. El exceso de la longitud sobre la anchura lo he multiplicado por sí mismo y el resultado por 9. Y este área es el área obtenida multiplicando la longitud por ella misma. ¿Cuáles son la longitud y la anchura?» Es evidente que aquí las palabras *longitud, anchura* y *área* son simplemente nombres cómodos para las dos incógnitas y su producto, respectivamente [1].

Hoy escribiríamos este problema simplemente como

$$xy = 10$$
$$9(x - y)^2 = x^2$$

La solución, dicho sea de paso, conduce a una ecuación de cuarto grado en x, en la que faltan los términos en x^3 y en x, de manera que en realidad es lo que nosotros llamamos una ecuación bicuadrada, que se puede resolver como una ecuación cuadrática en x^2, y así lo hicieron los antiguos babilonios.

[1] Pueden verse muchos ejemplos de problemas algebraicos en el libro de Van der Waerden *Science Awakening*, Noordhoff, 1954, págs. 65-73.

También aparecen problemas que conducen a raíces cúbicas; uno de estos problemas, formulado en simbolismo moderno, es el siguiente:

$$12x = z, \qquad y = x, \qquad xyz = V,$$

donde V es un volumen dado. Aquí, para calcular x tenemos que extraer una raíz cúbica; para ello los babilonios calculaban dicha raíz a partir de las tablas de cubos y raíces cúbicas que hemos mencionado más arriba. Aparecen también problemas de interés compuesto en los que se pide calcular el valor de una incógnita que figura en un exponente.

En realidad los babilonios utilizaron a veces símbolos especiales para las incógnitas, pero este simbolismo pasó inadvertido. En algunos problemas aparecen dos palabras sumerias especiales (un poco modificadas por terminaciones acadias) para representar dos incógnitas que son una inversa de la otra. Además, se utilizaban de hecho los antiguos pictogramas sumerios para estas palabras, y como ya no se usaban tales pictogramas en el lenguaje usual, el efecto era el mismo que si se utilizasen símbolos especiales para representar las incógnitas. Estos símbolos se usaron repetidamente y se pueden identificar fácilmente, incluso sin saber cómo se pronunciaban en acadio.

En la resolución de los problemas algebraicos solamente se iban explicando las etapas necesarias para llegar a la solución. Por ejemplo: «eleva al cuadrado 10, lo que da 100; resta 100 de 1.000, lo que da 900», etc. Dado que no aparece razón alguna que justifique cada etapa, lo único que podemos hacer nosotros es inferir cómo sabían lo que había que hacer.

En algunos problemas concretos sumaban los babilonios progresiones aritméticas y geométricas; nos encontramos, por ejemplo, en nuestra notación, con la suma

$$1 + 2 + 4 + \ldots + 2^9 = 2^9 + (2^9 - 1) = 2^{10} - 1.$$

También aparece la suma de los cuadrados de los números enteros del 1 al 10, como si se hubiera calculado aplicando la fórmula

$$1^2 + 2^2 + \ldots + n^2 = \left(1 \cdot \frac{1}{3} + n \cdot \frac{2}{3}\right)(1 + 2 + 3 + \ldots + n).$$

Sin embargo, los casos concretos que aparecen en los textos no vienen acompañados de demostración alguna.

El álgebra babilónica incluía también algo de lo que nosotros llamamos teoría de números. Así, aparecen calculadas muchas ternas pitagóricas, probablemente aplicando la regla correcta, es decir, que si

$$x = p^2 - q^2$$
$$y = 2pq$$
$$z = p^2 + q^2$$

entonces

$$x^2 + y^2 = z^2.$$

Y también resolvieron la ecuación $x^2 + y^2 = 2z^2$ para números enteros.

6. La geometría babilónica

El papel de la geometría en Babilonia fue prácticamente insignificante, no llegando a constituir una rama independiente de la matemática. Los problemas sobre divisiones de campos o sobre tamaños de ladrillos necesarios para alguna construcción se convertían inmediatamente en problemas algebraicos. Algunos cálculos de áreas y volúmenes se daban siguiendo ciertas reglas o fórmulas; sin embargo, las figuras que ilustran los problemas geométricos aparecen dibujadas toscamente y las fórmulas utilizadas a menudo son incorrectas. En los cálculos babilónicos de áreas, por ejemplo, no puede decirse con seguridad si los triángulos son rectángulos o si los cuadriláteros son cuadrados y, por lo tanto, si las fórmulas aplicadas son correctas o no para las figuras en cuestión. Sin embargo, ya se conocían la relación pitagórica, la semejanza de triángulos y la proporcionalidad de los lados correspondientes en triángulos semejantes. Al parecer, el área del círculo se calculaba siguiendo la regla $A = c^2/12$, donde c es la longitud de la circunferencia; esta regla equivale, evidentemente, a utilizar 3 como valor de π. Sin embargo, otro de sus resultados, en el que se da la relación entre el perímetro de un hexágono regular y su circunferencia circunscrita, supone adoptar un valor de 3 1/8 para π.

Se sabía calcular, unos correcta y otros incorrectamente, algunos volúmenes que se presentaban al resolver problemas físicos concretos.

Aparte de algunos hechos especiales, tales como el cálculo del radio de la circunferencia circunscrita a un triángulo isósceles dado, la geometría babilónica venía a reducirse a una colección de reglas para el cálculo de áreas de figuras planas sencillas, incluyendo los polígonos regulares, y de los volúmenes de cuerpos sólidos también sencillos. La geometría no se estudió nunca en sí y por sí misma, sino siempre en conexión con problemas prácticos.

7. Aplicaciones de la matemática en Babilonia

A pesar de su limitada extensión, la matemática entraba en muchos aspectos de la vida de los babilonios. Babilonia era un cruce de importantes rutas comerciales, y los babilonios utilizaron sus conocimientos de aritmética y de álgebra elemental aplicados a longitudes y pesos, a intercambios de moneda y mercancías, al cálculo de interés simple y compuesto, de los impuestos y de las porciones de una cosecha a pagar al granjero, al templo y al Estado, mientras que los problemas de herencias y divisiones de campos conducían a problemas algebraicos. La mayoría de los textos cuneiformes que tratan de matemáticas (excluyendo las tablas y textos de ejercicios) se refieren a problemas económicos. No hay duda, pues, de la influencia de la economía en el desarrollo de la aritmética del período más antiguo.

La construcción de canales, presas y otros proyectos de riego exigía cálculos, y el uso de ladrillos planteaba numerosos problemas numéricos y geométricos. Otros cálculos útiles eran los de volúmenes de graneros y edificios, y los de áreas de campos. La estrecha relación entre la matemática babilónica y los problemas prácticos aparece tipificada en lo siguiente: se trata de excavar un canal de sección trapezoidal y de dimensiones dadas. Se conoce también lo que puede cavar un hombre en un día, así como la suma del número de hombres empleados y los días que han de trabajar. El problema consiste en calcular el número de hombres y el número de días de trabajo.

Dado que la conexión entre matemática y astronomía se hizo esencial desde la época de los griegos en adelante, es interesante saber qué conocían los babilonios sobre astronomía. De la astronomía sumeria no sabemos nada, y durante el período acadio la astronomía fue cualitativa y rudimentaria; indudablemente el desarrollo de la

matemática precedió al desarrollo de cualquier tipo importante de astronomía. Durante el período asirio (hacia el 700 a. C.), la astronomía comenzó a incluir descripciones matemáticas de los fenómenos y un registro sistemático de los datos de observación. El uso de la matemática aumentó sustancialmente durante los tres últimos siglos antes de nuestra era, dedicándose de manera especial al estudio de los movimientos lunares y planetarios; de hecho, la mayor parte de los textos astronómicos data de este período seléucida. Estos textos pueden clasificarse en dos grupos: efemérides planetarias y tablas de posiciones de los cuerpos celestes en diversas épocas. Hay indicaciones de cómo calcular las efemérides.

La aritmética que hay detrás de las observaciones lunares y solares muestra que los babilonios calculaban las diferencias primeras y segundas de los datos sucesivos, observaban la constancia de esas diferencias primeras o segundas y extrapolaban o interpolaban para otros datos. Su manera de proceder equivalía a utilizar el hecho de que los datos pueden ajustarse mediante funciones polinómicas, lo que les permitía predecir la posición diaria de los planetas. Conocían con cierta exactitud los períodos de los planetas y utilizaban los eclipses como base de cálculo. No hubo, sin embargo, ningún esquema geométrico del movimiento lunar o planetario en la astronomía babilónica.

Los babilonios del período seléucida disponían ya de extensas tablas sobre los movimientos del Sol y de la Luna, que les daban velocidades y posiciones variables. También aparecían en las tablas, o se podían obtener fácilmente de ellas, conjunciones especiales y eclipses del Sol y de la Luna; así pues, los astrónomos podrían predecir las lunas nuevas y los eclipses dentro de un margen de pocos minutos. Sus datos indican que conocían la longitud del año solar o tropical (o año de las estaciones) como $12 + 22/60 + 8/60^2$ meses (de Luna nueva a Luna nueva), y la longitud del año sidéreo (el tiempo que emplea el Sol en volver a la misma posición relativa a las estrellas) con menos de 4 1/2 minutos de margen.

Las constelaciones que dan sus nombres a los doce signos del Zodiaco ya se conocían antes, pero el Zodiaco mismo aparece por primera vez en un texto del año 419 a. C. Cada sector del Zodiaco abarcaba un arco de 30 grados y las posiciones de los planetas en el cielo se fijaban con respecto a las estrellas y también por su posición en el Zodiaco.

La astronomía servía para fines muy diversos. Para empezar, era

necesaria para hacer el calendario, que venía determinado por las posiciones del Sol, la Luna y las estrellas. El año, el mes y el día son cantidades astronómicas que hay que calcular con exactitud para conocer la época de la siembra y las fiestas religiosas, por ejemplo. En Babilonia, debido en parte a la conexión del calendario con las fiestas y ceremonias religiosas, y en parte a que los cuerpos celestes eran considerados como dioses, eran los sacerdotes los encargados de llevar el calendario.

Este calendario era básicamente lunar; el mes comenzaba con la primera aparición del cuarto creciente después del oscurecimiento total de la Luna o Luna nueva, mientras que el día comenzaba por la tarde de la primera aparición del cuarto creciente y duraba de puesta del Sol a puesta del Sol. El calendario lunar es difícil de mantener porque, aunque conviene que el mes contenga un número entero de días, los meses lunares, calculados a base del tiempo entre dos conjunciones sucesivas del Sol y de la Luna (es decir, de Luna nueva a Luna nueva), varía de 29 a 30 días. Por lo tanto, se plantea el problema de decidir qué meses han de tener 29 y cuáles 30. Otro problema, más importante aún, es el de poner de acuerdo el calendario lunar con las estaciones. La solución es muy complicada porque depende de las trayectorias y velocidades del Sol y de la Luna. El calendario lunar contenía meses extra intercalados entre los normales, de manera que 7 de éstos cada 19 años venían a mantener *aproximadamente* de acuerdo el calendario lunar con el año solar, de manera que 235 meses lunares equivalían a 19 años solares. Se calculaba sistemáticamente el solsticio de verano, y tanto el solsticio de invierno como los equinoccios se colocaban a intervalos iguales. Este calendario fue utilizado por los judíos, los griegos y los romanos hasta el 45 a. C., en que se adoptó el calendario llamado Juliano.

La división del círculo en 360 grados tuvo su origen en la astronomía babilónica de los últimos siglos anteriores a la era cristiana, y no parece haber tenido nada que ver con la utilización anterior de la base 60; sin embargo, la base 60 sí se usó para dividir el grado y el minuto en 60 partes cada uno, y el astrónomo Ptolomeo (siglo II d. C.) siguió a los babilonios en esta práctica.

Estrechamente relacionada con la astronomía estuvo la astrología. En Babilonia, como en tantas civilizaciones antiguas, los cuerpos celestes se consideraron como dioses y, por lo tanto, se suponía que tenían influencia e incluso control sobre los asuntos de los hombres. Teniendo en cuenta la importancia del Sol para la luz, el calor y el

crecimiento de las plantas, el terror inspirado por sus eclipses, y muchos fenómenos estacionales como el apareamiento de los animales, resulta fácil entender la creencia de que los cuerpos celestes afectan incluso a los acontecimientos diarios en la vida del hombre.

Los sistemas de predicción seudocientíficos en las antiguas civilizaciones no siempre tuvieron que ver con la astronomía; los números mismos tenían presuntas propiedades místicas y podían utilizarse también para hacer predicciones. Podemos encontrar algunos usos babilónicos en el Libro de Daniel y en los escritos de los profetas del Antiguo y Nuevo Testamento. La «ciencia» hebrea de la *gematría* (una forma de misticismo cabalístico) se basaba en el hecho de que cada letra del alfabeto tenía un valor numérico determinado, porque los hebreos usaban las letras para representar números. Si la suma de los valores numéricos de las letras en dos palabras era la misma, se deducía una importante conexión entre las dos ideas o personas o sucesos representados por esas dos palabras. En la profecía de Isaías (21:8), el león proclama la caída de Babilonia, debido a que las letras en la palabra hebrea para león y para Babilonia suman lo mismo.

8. Evaluación global de la matemática babilónica

La utilización por parte de los babilonios de términos y símbolos especiales para las incógnitas, el uso de algunos símbolos operativos y su solución de algunos tipos de ecuaciones con una o más incognitas, especialmente las ecuaciones cuadráticas, constituye el punto de partida del álgebra. El desarrollo de un método sistemático para escribir números enteros y fracciones les permitió disponer de una aritmética bastante avanzada y utilizarla en muchas situaciones prácticas, especialmente en astronomía. Podríamos decir que alcanzaron un tipo de habilidad numérica y algebraica para resolver ecuaciones especiales de grado más alto, pero, consideradas globalmente, su aritmética y su álgebra fueron muy elementales. A pesar de que trabajaban con números y problemas concretos, mostraron un cierto grado de abstracción matemática al reconocer que algunos métodos eran propios de determinadas clases de ecuaciones.

También se plantea la cuestión de hasta qué punto utilizaron los babilonios la idea de demostración en matemáticas. Efectivamente, resolvieron, mediante procedimientos sistemáticos correctos, ecuaciones bastante complicadas y, sin embargo, se limitaban a dar instruc-

ciones verbales de los cálculos a realizar, sin justificarlos de ninguna manera. Es casi seguro que los procesos aritméticos y algebraicos y las reglas geométricas que utilizaron fueran el resultado final de la evidencia física misma, acompañada del método de ensayo y error; para los babilonios resultaba justificación suficiente para seguir utilizando dichos procesos el que funcionasen aceptablemente bien. En resumen, en la matemática babilónica no se encuentra ni el concepto de demostración, ni la idea de una estructura lógica basada en principios que merecieran aceptación por un motivo u otro, ni la consideración de cuestiones tales como las de bajo qué condiciones pueden existir soluciones de los problemas.

Bibliografía

Bell, E. T.: *The Development of Mathematics*, 2.ª ed., McGraw-Hill, capítulos 1-2.

Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1894, vol. 1, cap. 1.

Chiera, E.: *They Wrote on Clay*, Chicago University Press, 1938.

Childe, V. Gordon: *Man Makes Himself*, New American Library, 1951, caps. 6-8.

Dantzig, Tobias: *Number: The Language of Science*, 4.ª ed., Macmillan, 1954, caps. 1-2.

Karpinski, Louis C.: *The History of Arithmetic*, Rand McNally, 1925.

Menninger, K.: *Number Words and Number Symbols: A Cultural History of Numbers*, Massachusetts Institute of Technology Press, 1969.

Neugebauer, Otto: *The Exact Sciences in Antiquity*, Princeton University Press, 1952, caps. 1-3 y 5.

—: *Vorgriechische Mathematik*, Julius Springer, 1934, caps. 1-3 y 5.

Sarton, George: *A History of Science*, Harvard Universy Press, 1952, vol. 1, cap. 3.

—: *The Study of the History of Mathematics and the History of Science*, Dover (reprint), 1954.

Smith, David Eugene: *History of Mathematics*, Dover (reprint), 1958, vol. 1, cap. 1.

Struik, Dirk J.: *A Concise History of Mathematics*, 3.ª ed., Dover, 1967, capítulos 1-2.

Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, caps. 2-3.

Capítulo 2
LA MATEMATICA EGIPCIA

La ciencia toda, incluida la lógica y la matemática, es función de la época; la totalidad de la ciencia, tanto en sus ideales como en sus logros.

E. H. MOORE

1. El marco histórico

Mientras que en Mesopotamia los pueblos que ejercieron el dominio sociopolítico del país a lo largo de su historia cambiaron con frecuencia, con el resultado de la aparición de nuevas influencias culturales, la civilización egipcia se desarrolló sin verse afectada prácticamente por influencias extranjeras. Desconocemos los orígenes de esta civilización, pero seguramente existía ya incluso antes del 4000 a. C. Egipto, como decía el historiador griego Herodoto, es un regalo del Nilo. Una vez al año, este río, que recoge sus aguas en el lejano sur del Africa central y de Abisinia, inunda casi todo el territorio que se extiende a lo largo de sus riberas, y deja fértiles depósitos de limo al retirarse. La mayor parte de la población vivía de cultivar estas tierras, y aún hoy lo sigue haciendo; el resto del país es un desierto.

Al principio hubo dos reinos, uno en el norte y otro en el sur de lo que es hoy Egipto, hasta que, en algún momento entre el 3500 y el 3000 a. C., el rey Menes unificó los llamados Alto y Bajo Egipto. A partir de ese momento, los grandes períodos de la historia egipcia se establecen cronológicamente en términos de las distintas dinastías reinantes, considerando a Menes como el fundador de la primera

dinastía. La culminación de la cultura egipcia se produjo en torno a la tercera dinastía (hacia el 2500 a. C.), durante la cual los faraones hicieron construir las grandes pirámides. La civilización egipcia siguió sus propios derroteros hasta que Alejandro Magno conquistó el país el año 332 a. C. En adelante, y hasta poco después del 600 d. C., tanto su historia como su matemática pertenecen ya a la cultura griega. Así pues, dejando aparte una invasión menor de los hiksos (1700-1600 a. C.), y algunos contactos con la civilización babilónica (que se deducen del descubrimiento en el valle del Nilo de las llamadas tablillas cuneiformes de Tell al-Amarna, de hacia el 1500 a. C.), la civilización egipcia fue una creación altamente original del pueblo que vivió durante esos tres milenios en el valle del Nilo.

Los antiguos egipcios desarrollaron sus propios sistemas de escritura. Uno de ellos, y el más antiguo, la escritura jeroglífica, era de tipo pictórico, es decir, que cada símbolo era el dibujo de algún objeto concreto. La escritura jeroglífica se utilizó mucho en los monumentos hasta comienzos de nuestra era, pero desde el 2500 a. C., aproximadamente, los egipcios usaron en la vida diaria, al escribir sobre papiro, la llamada escritura hierática. Este sistema utilizaba símbolos convencionales que en principio habían sido meras simplificaciones de los símbolos jeroglíficos por un proceso de estilización. La escritura hierática es silábica; cada sílaba venía representada por un ideograma, y una palabra completa por una colección de ideogramas. El significado de cada palabra no tiene, en general, nada que ver con el de cada ideograma por separado.

La escritura usual se hacía con tinta negra o roja sobre papiro. Las hojas de papiro se producían cortando en tiras delgadas el tallo de la planta del mismo nombre, pegando estas tiras en dos capas cruzadas, prensándolas y encolándolas. Debido a que el papiro, al secarse excesivamente, se resquebraja, nos han llegado pocos documentos del antiguo Egipto, dejando aparte las inscripciones jeroglíficas sobre piedra, abundantes, pero que transmiten escasa información interesante.

Los documentos matemáticos más importantes que han sobrevivido son dos papiros bastante extensos: el papiro de Moscú, que se conserva en un museo de la capital rusa, y el papiro Rhind, descubierto en 1858 por el anticuario escocés A. Henry Rhind y ahora en el British Museum. El papiro Rhind también se conoce como papiro de Ahmes, por el nombre de su autor, que comienza con las siguientes palabras: «Cálculo Exacto para Entrar en Conocimiento de Todas las

Cosas existentes y de Todos los Oscuros Secretos y Misterios.» Ambos papiros datan de hacia el 1700 a. C. También hay algunos fragmentos de otros papiros escritos en la misma época y posteriores. Estos papiros de tipo matemático fueron redactados por escribas que eran funcionarios del Estado egipcio o administradores de los templos.

Lo que contienen los papiros son problemas y sus soluciones, 85 en el papiro Rhind y 25 en el papiro de Moscú. Es posible que tales problemas se les presentasen a los escribas en su trabajo, y se esperaba que los supiesen resolver, pero lo más probable es que los problemas que figuran en los dos papiros más importantes tuvieran una intención pedagógica, como ejemplos más o menos artificiales de problemas típicos y sus soluciones. A pesar de que estos papiros datan, como hemos dicho, de hacia 1770 a. C., las matemáticas que aparecen en ellos probablemente las conocían ya los egipcios en fecha tan remota como el 3500 a. C., y poco fue lo que se añadió desde esa época hasta la conquista griega.

2. La aritmética

Los símbolos numéricos jeroglíficos que utilizaron los egipcios fueron | para el 1, ∩ para 10, 𓍢 o 𓍢 para 100, 𓆼 para 1.000, 𓂭 para 10.000, 𓆐 100.000 y 𓁨 para 1.000.000. Estos símbolos se combinaban para formar números intermedios, siendo la dirección de la escritura de derecha a izquierda, de manera que ||||∩∩, por ejemplo, representaba 24. Se trata, pues, de un sistema de escritura numérica que usa la base 10, pero no es posicional, sino aditivo.

En escritura hierática egipcia, los símbolos para los diez primeros números naturales son los siguientes:

1 2 3 4 5 6 7 8 9 10

La aritmética egipcia fue esencialmente aditiva; para las sumas y restas usuales se limitaban a combinar o a cancelar los diferentes

símbolos hasta llegar al resultado concreto. La multiplicación y la división también se reducían en último término a procesos aditivos, pero el cálculo era un poco más complicado. Para calcular 12 por 12, por ejemplo, los egipcios hacían lo siguiente:

1	12
2	24
4	48
8	96.

Cada línea se obtiene de la anterior por duplicación, y como 4 + 8 = 12 y 4 · 12 = 48 y 8 · 12 = 96, sumando 48 y 96 se obtiene el valor de 12 por 12. Como se ve, este proceso es bastante distinto del usual de multiplicar por 10, luego por 2, y sumar. La multiplicación por 10 se efectuaba a veces sustituyendo los símbolos de las decenas por símbolos para 100, etc.

Particularmente interesante resultá el método utilizado por los egipcios para dividir un número por otro. Por ejemplo, para dividir 19 por 8 procedían de la manera siguiente:

1	8
2	16
1/2	4
1/4	2
1/8	1

y, por lo tanto, la respuesta era 2 + 1/4 + 1/8. La idea consiste simplemente en tomar el número de ochos y de partes de 8 que sumen 19.

El método de representación de las fracciones en el sistema numérico egipcio era mucho más complicado que el nuestro. El símbolo ⬭, que se pronuncia *ro* y que originariamente representaba 1/320 de un bushel, terminó por representar la idea de fracción; en escritura hierática este símbolo oval se sustituyó por un punto. El símbolo ⬭ o el punto se solía escribir encima del número para indicar la correspondiente fracción unitaria. Así, por ejemplo, en escritura jeroglífica

$$\overset{⬭}{\text{||| ||}} = \frac{1}{5}, \qquad \overset{⬭}{\cap} = \frac{1}{10}, \qquad \overset{⬭}{\cap}\,\text{||| ||} = \frac{1}{15}.$$

Los egipcios disponían de unos pocos símbolos especiales para algunas fracciones muy concretas. Así, el jeroglífico ⊂= representaba 1/2; ⫪ , 2/3; y ×, 1/4.

Aparte de esas pocas especiales, todas las demás fracciones se descomponían en lo que llamamos fracciones unitarias. Así, por ejemplo, Ahmes escribe 2/5 como 1/3 + 1/15, donde el signo más no aparece pero se sobreentiende. El papiro Rhind contiene al principio una tabla en la que se expresan las fracciones de numerador 2 y de denominador impar entre 5 y 101, como sumas de fracciones unitarias. Por medio de esta tabla una fracción tal como la 7/29, que para Ahmes significa 7 dividido por 29, podría expresarse como suma de fracciones unitarias: dado que 7 = 2 + 2 + 2 + 1, Ahmes procede a convertir cada 2/29 en una suma de fracciones unitarias; combinando estos resultados y modificándolos un poco llega a una suma de fracciones unitarias, todas de distinto denominador, que dan la expresión final para 7/29 en la forma

$$\frac{1}{6}+\frac{1}{24}+\frac{1}{58}+\frac{1}{87}+\frac{1}{232}.$$

Es fácil comprobar que 7/29 también puede expresarse como 1/5 + 1/29 + 1/145, pero como la tabla de $2/n$ de Ahmes conduce a la expresión anterior, es ésta la que se usa. La expresión de una de nuestras fracciones a/b como suma de fracciones unitarias se practicó de manera sistemática en Egipto siguiendo métodos y reglas elaborados a lo largo de siglos desde una remota antigüedad. Los egipcios efectuaban las cuatro operaciones aritméticas con fracciones utilizando las fracciones unitarias. Los frecuentes y complicados cálculos con fracciones fueron sin duda una de las razones de que los egipcios no llegaran a desarrollar nunca una aritmética ni un álgebra avanzadas.

La naturaleza de los números irracionales tampoco llegó a reconocerse en la aritmética egipcia, al igual que no lo había sido en la babilónica. Las raíces cuadradas sencillas que aparecían en los problemas aritméticos o algebraicos se podían expresar, y se expresaron, en términos de números enteros y de fracciones.

3. Algebra y geometría

Los papiros que nos han llegado contienen también soluciones de problemas con una incógnita, que vienen a ser equivalentes a nuestra resolución de ecuaciones lineales. Sin embargo, los procesos seguidos eran puramente *aritméticos* y no constituían, para los egipcios, un tema distinto, como podía ser la resolución de ecuaciones. Estos problemas aparecen formulados verbalmente, como todos, con unas someras instrucciones para obtener la solución, sin explicación alguna de por qué se usan tales procedimientos ni de por qué funcionan bien. Por ejemplo, el problema 31 del papiro de Ahmes, traducido literalmente, dice: «Una cantidad; sus 2/3, su 1/2, su 1/7, su totalidad asciende a 33.» Esto para nosotros significa:

$$\frac{2x}{3} + \frac{x}{2} + \frac{x}{7} + x = 33.$$

La solución viene dada en este caso en términos de simples operaciones aritméticas del tipo egipcio, que ya hemos visto.

El problema 63 del mismo papiro dice lo siguiente: «Instrucciones para dividir 700 hogazas de pan entre 4 personas, 2/3 para el primero, 1/2 para el segundo, 1/3 para el tercero, 1/4 para el cuart.». Esto significa para nosotros resolver la ecuación

$$\frac{2x}{3} + \frac{x}{2} + \frac{x}{3} + \frac{x}{4} = 700.$$

La solución dada por Ahmes es la siguiente: «Suma 2/3, 1/2, 1/3, 1/4; esto da $1\,\frac{1}{2}\,\frac{1}{4}$. Divide 1 por $1\,\frac{1}{2}\,\frac{1}{4}$; esto da $\frac{1}{2}\,\frac{1}{14}$. Ahora calcula el $\frac{1}{2}\,\frac{1}{14}$ de 700; esto da 400.»

En algunos casos Ahmes utiliza en su solución la llamada «regula falsi», o «regla de la falsa posición». Así, para calcular cinco números en progresión aritmética, sujetos a una condición extra y tales que su suma sea 100, elige Ahmes la diferencia d de la progresión de manera que sea igual a 5 1/2 veces el término menor, y toma tal término menor igual a 1, con lo que obtiene la progresión 1, 6 1/2, 12, 17 1/2 y 23. Pero estos números sólo suman 60, mientras que lo que debían

sumar era 100. Ahmes multiplica entonces cada uno de los términos por 5/3 = 100/60.

El único tipo de ecuación de segundo grado que aparece es el más sencillo, $ax^2 = b$; incluso donde aparecen dos incógnita, el problema es del tipo

$$x^2 + y^2 = 100; \quad y = \frac{3}{4}x,$$

de manera que eliminando la y, la ecuación en x se reduce efectivamente al primer tipo. También nos encontramos en los papiros algunos problemas concretos en los que aparecen progresiones aritméticas y geométricas. En todos estos problemas no resulta muy difícil inferir reglas generales a partir de las soluciones.

Este álgebra egipcia tan restringida no utilizaba prácticamente ningún simbolismo. En el papiro de Ahmes las operaciones de sumar y restar aparecen representadas por un dibujo esquemático de las piernas de una persona que se acerca y que se aleja, respectivamente, es decir y , y el símbolo se utiliza para representar una raíz cuadrada.

Y ¿qué se puede decir de la geometría egipcia? En realidad, lo primero que hay que señalar es que los egipcios no establecían ninguna separación entre aritmética y geometría, y en los papiros nos encontramos con problemas de los dos tipos mezclados. Al igual que los babilonios, los egipcios consideraban la geometría como una herramienta práctica. Uno se limitaba a aplicar la aritmética y el álgebra a problemas de áreas, volúmenes y otras situaciones geométricas. Herodoto nos dice que la geometría egipcia tuvo su origen en la necesidad que provocaba la crecida anual del Nilo de volver a trazar las lindes de los terrenos cultivados por los agricultores. Sin embargo, los babilonios desarrollaron una geometría parecida sin tal necesidad. Los egipcios disponían de recetas para el cálculo de áreas de rectángulos, triángulos y trapezoides; en el caso del área de un triángulo, aunque multiplicaban un número por la mitad de otro, no podemos estar seguros de que el método sea correcto, porque no tenemos la seguridad de que las palabras utilizadas representen las longitudes de la base y la altura o simplemente dos lados. Además, las figuras están tan mal dibujadas en los papiros que a veces no se puede saber exactamente qué área o volumen se está calculando. Su cálculo del

área del círculo, sorprendentemente bueno, usa la fórmula $A = (8d/9)^2$ donde d es el diámetro, lo que supone utilizar 3,1605 como valor de π.

Un ejemplo puede ilustrarnos bien la «exactitud» de las fórmulas egipcias para áreas. En los muros de un templo de Edfu aparece una lista de campos, presumiblemente regalos al templo; estos campos solían tener cuatro lados, que representaremos por a, b, c, d, donde a, b y c, d son las parejas de lados opuestos. Las inscripciones dan las áreas de estos campos siguiendo la regla $\frac{(a + b)}{2} \cdot \frac{(c + d)}{2}$. Pero algunos campos son triangulares y en ese caso se dice que d es nada y el cálculo se transforma en el de $\frac{(a + b)}{2} \cdot \frac{c}{2}$. Incluso para cuadriláteros constituye esta regla una aproximación muy grosera.

Los egipcios también tenían reglas para el volumen de un cubo, un paralelepípedo, un cilindro y otras figuras sencillas, algunas de ellas correctas y otras sólo aproximadas. Los papiros dan como volumen de un tronco de cono, que representa probablemente una clepsidra (o reloj de agua), el siguiente:

$$V = \frac{h}{12}\left[\frac{3}{2}(D + d)\right]^2,$$

donde h es la altura y $(D + d)/2$ es la circunferencia media. Esta fórmula supone utilizar 3 como valor de π.

La regla más sorprendente quizás de la geometría egipcia es la del volumen de un tronco de pirámide de base cuadrada que, escrita en notación moderna es

$$V = \frac{h}{3}(a^2 + ab + b^2),$$

donde h es la altura y a y b las aristas básicas. La fórmula es sorprendente porque es correcta y porque aparece expresada de manera simétrica, aunque no, desde luego, en nuestra notación, sino que viene dada para números concretos solamente. Sin embargo, no sabemos ni siquiera si la pirámide es de base cuadrada o no por lo defectuoso de la figura dibujada en el papiro.

Tampoco sabemos si los egipcios reconocieron el teorema de

Pitágoras. Sí sabemos que había agrimensores o «tensadores de la cuerda», pero la historia de que utilizaban una cuerda anudada a intervalos iguales para dividir la longitud total en partes de longitudes 3, 4 y 5, que podían usar para formar un triángulo rectángulo, no aparece confirmada en ningún documento.

Las reglas formuladas no aparecen expresadas en símbolos, naturalmente. Los egipcios enunciaban los problemas verbalmente, y su procedimiento para resolverlos era esencialmente lo que nosotros hacemos cuando calculamos siguiendo una fórmula. Así, por ejemplo, una traducción casi literal del problema geométrico de calcular el volumen de un tronco de pirámide es la siguiente: «Si te dicen: una pirámide truncada de 6 como altura vertical por 4 en la base por 2 en el extremo superior. Tienes que cuadrar este 4, resultado 16. Tienes que doblarlo, resultado 8. Tienes que cuadrar 2, resultado 4. Tienes que sumar el 16, el 8 y el 4, resultado 28. Tienes que tomar un tercio de 6, resultado 2. Tienes que tomar dos veces el 28, resultado 56. Ves, es 56. Lo has hecho correctamente.»

¿Conocían los egipcios demostraciones o justificaciones de sus procedimientos y recetas? Algunos creen que el papiro de Ahmes fue escrito en el estilo de un libro de texto para estudiantes de la época y que, por lo tanto, aunque Ahmes no formule ninguna regla o principio general para resolver diferentes tipos de ecuaciones, es muy probable que las conociera, pero quería que el estudiante las formulara por sí mismo o bien tuviera cerca un maestro que lo hiciera por él. Bajo este punto de vista, el papiro de Ahmes resulta un texto de aritmética bastante avanzado. Otros creen que se trata del cuaderno de notas de un alumno. En cualquier caso, los papiros registran casi con toda seguridad los tipos de problemas que debían resolver los escribas en asuntos de negocios y administrativos, y que los métodos de resolución eran simplemente reglas prácticas conocidas por experiencia en ese trabajo. Nadie cree seriamente que los egipcios dispusieran de una estructura deductiva, basada en axiomas, que justificara la corrección de sus reglas.

4. Aplicaciones de la matemática egipcia

Los egipcios utilizaron la matemática en la administración de los asuntos del Estado y de los templos, en el cálculo de salarios pagados a los trabajadores, en el cálculo de volúmenes de graneros y áreas de

campos, en el cobro de impuestos estimados según el área de la tierra, en la conversión de un sistema de medidas a otro y en el cálculo del número de ladrillos necesario para la construcción de edificios o rampas. Los papiros contienen también problemas relativos a la cantidad de grano necesario para producir cantidades dadas de cerveza, o la cantidad de grano de una calidad necesario para obtener el mismo resultado que con grano de otra calidad, cuya «fuerza» relativa al primero fuera conocida.

Como en Babilonia, se hacía un importante uso de la matemática en astronomía, cosa que data de la primera dinastía. Para los egipcios los conocimientos astronómicos eran esenciales, por lo siguiente. El Nilo es el elemento esencial de la vida en Egipto, cuyos habitantes viven de cultivar las tierras que el Nilo cubre de rico mantillo en su desbordamiento anual. Sin embargo, el egipcio tenía que estar bien preparado para los aspectos peligrosos de la inundación; casa, herramientas y ganado tenían que ser retirados temporalmente de la zona y hacer los preparativos para sembrar inmediatamente después. Por lo tanto, era necesario predecir la llegada de la inundación, cosa que se hacía por los fenómenos astronómicos que la precedían.

La astronomía hizo posible también el calendario. Aparte de la necesidad del calendario para el comercio, estaba la necesidad de establecer las fiestas religiosas, puesto que se creía esencial, para asegurar la benevolencia de los dioses, que las fiestas se celebraran en el momento debido. Y, lo mismo que en Babilonia una vez más, la tarea de llevar el calendario correspondió en gran parte a los sacerdotes.

Los egipcios llegaron a calcular la longitud del año solar observando la estrella Sirio. Un día del verano se hacía visible esta estrella en el horizonte exactamente antes de la salida del sol, mientras que en días sucesivos permanecía visible durante más tiempo antes de que la luz del sol la extinguiese. El momento en que era visible justo antes de la salida del sol recibía el nombre de salida heliacal de Sirio, y el intervalo entre dos de ellas consecutivas era de, aproximadamente, 365 1/4 días. Así pues, los egipcios adoptaron (se supone que el 4241 a. C.) un calendario civil con un año de 365 días. La concentración en Sirio se debió indudablemente al hecho de que las aguas del Nilo comenzaban a subir aproximadamente ese día, que se eligió como primer día del año.

El año de 365 días se dividió en 12 meses de 30 días, más cinco días extras al final. Como los egipcios no intercalaron el día adicional cada

cuatro años, el calendario civil iba retrasándose poco a poco con respecto a las estaciones, y al cabo de 1460 años volvía a la situación inicial; a este intervalo se le llama ciclo Sótico, del nombre egipcio para Sirio. No sabemos con seguridad si los egipcios conocieron este ciclo. Su calendario fue adoptado por Julio César el 45 a. C., pero transformado en un año de 365 1/4 días por consejo del griego alejandrino Sosígenes. A pesar de que la determinación del año por los egipcios y su calendario fueron contribuciones valiosas, esto no condujo a una astronomía bien desarrollada, sino que fue, de hecho, rudimentaria y muy inferior a la babilónica.

Los egipcios combinaron sus conocimientos de astronomía y de geometría para construir sus templos, de manera que en ciertos días del año el sol incidiera sobre ellos de una manera especial. Por ejemplo, algunos fueron construidos de manera que el día más largo del año el sol penetraba directamente hasta el fondo del santuario e iluminaba la efigie del dios sobre el altar. Esta orientación de los templos la encontramos también a veces en Babilonia y en Grecia. Las pirámides se orientaban igualmente en direcciones especiales con respecto al cielo, y la Esfinge mira hacia el Este. Aunque los detalles de la construcción de estos monumentos no nos interesan ahora, vale la pena observar que las pirámides representan otra aplicación de la geometría egipcia. Son tumbas de faraones, como se sabe, y, dado que los egipcios creían en la inmortalidad, suponían que una tumba bien construida era esencial para la otra vida; de hecho, en cada pirámide se instalaba una residencia completa para el rey y la reina, y se ponía especial cuidado en construir sus bases de forma correcta, y las dimensiones relativas de la base y la altura eran importantes. Sin embargo, no hay que exagerar la complejidad o profundidad de las ideas puestas en juego; la matemática egipcia fue simple y rudimentaria, y no incluía principios profundos, contrariamente a lo que suele afirmarse.

5. Resumen

Revisemos sumariamente la situación de la matemática antes de que los griegos entren en escena. En las civilizaciones babilónica y egipcia nos encontramos con una aritmética de números enteros y fracciones, incluida la notación posicional, los comienzos del álgebra y algunas fórmulas empíricas en geometría. Casi no hay simbolismo

apenas algún pensamiento consciente sobre abstracciones, ninguna formulación metodológica general y ninguna idea de demostración o incluso de razonamiento plausible que pudiera convencer a alguien de la corrección de un procedimiento o fórmula. No hubo, de hecho, ninguna concepción de ciencia teórica de ningún tipo.

Aparte de algunos resultados ocasionales en Babilonia, en ambas civilizaciones la matemática no se consideró una disciplina independiente digna de cultivarse por sí misma. Se trataba de una herramienta en forma de reglas simples y desconexas que respondían a problemas de la vida diaria, aunque ciertamente nada se hizo en matemáticas que alterase o afectase la forma de vida. A pesar de que la matemática babilónica fue más avanzada que la egipcia, casi lo mejor que se puede decir de ambas es que mostraron cierto vigor, si no rigor, y más perseverancia que brillantez.

Toda evaluación implica usar algún tipo de criterio. Puede resultar un tanto injusto, pero es natural comparar las dos civilizaciones con la griega que las sucedió. Con esta medida, los egipcios y los babilonios se nos presentan como rudos albañiles, mientras los griegos serían magníficos arquitectos. Pueden encontrarse descripciones más favorables, incluso elogiosas, de los logros de egipcios y babilonios, pero suelen estar hechas por especialistas en estas culturas, que se convierten, inconscientemente quizás, en devotos admiradores de su propio campo de interés.

Bibliografía

Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.
Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1894, vol. 1, cap. 3.
Chace, A. B., *et al.*, eds.: *The Rhind Mathematical Papyrus*, 2 vols., Mathematical Association of America, 1927-29.
Childe, V. Gordon: *Man Makes Himself*, New American Library, 1951.
Karpinski, Louis C.: *The History of Arithmetic*, Rand McNally, 1925.
Neugebauer, O.: *The Exact Sciences in Antiquity*, Princeton University Press, 1952, cap. 4.
—: *Vorgriechische Mathematik*, Julius Springer, 1934.
Sarton, George: *A History of Science*, Harvard University Press, 1952, vol. 1, cap. 2.
Smith, David Eugene: *History of Mathematics*, Dover (reprint), 1958, vol. 1, cap. 2; vol. 2, caps. 2 y 4.
Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, cap. 1.

Capítulo 3
LOS ORIGENES DE LA MATEMATICA CLASICA GRIEGA

> Así es, pues, la matemática: te recuerda la forma invisible del alma; da vida a sus propios descubrimientos; despiera la mente y purifica el intelecto; arroja luz sobre nuestras ideas intrínsecas y anula el olvido y la ignorancia que nos corresponden por nacimiento.
>
> PROCLO

1. El marco histórico

En la historia de la civilización los griegos alcanzaron una posición preeminente, y en la historia de la matemática su época fue una de las más brillantes. A pesar de que tomaron muchos elementos prestados de las civilizaciones vecinas, los griegos edificaron una civilización y una cultura originales, de las más impresionantes de toda la historia de la humanidad, la que más ha influido en el desarrollo de la cultura occidental moderna, y que fue decisiva en la fundamentación de la matemática tal como la entendemos hoy. Uno de los grandes problemas de la historia de la cultura es el de dar cuenta de la brillantez y de la creatividad de los antiguos griegos.

Aunque nuestro conocimiento de los orígenes de su historia está sujeto, evidentemente, a revisiones y clarificaciones según vayan avanzando las investigaciones arqueológicas, tenemos motivos para creer, sobre la base de la *Ilíada* y la *Odisea* de Homero, del desciframiento de las antiguas lenguas y escrituras, y de las mismas excavaciones arqueológicas, que la civilización griega se remonta hacia el 2800 a. C. Los griegos se instalaron en Asia Menor, que pudo haber sido su lugar de origen, en el territorio continental europeo que constituye la

Grecia moderna, y en el sur de Italia, Sicilia, Creta, Rodas, Delos y el norte de Africa. Hacia el 775 a. C., los griegos sustituyeron varios sistemas de escritura jeroglífica que utilizaban por la escritura alfabética fenicia (que también utilizaban ya los hebreos). Con la adopción del alfabeto, los griegos se convirtieron en un pueblo más letrado y mucho más capaz de registrar tanto su historia como sus ideas.

Con el establecimiento definitivo de los griegos en estos territorios, entraron en contacto comercial y cultural con los egipcios y los babilonios. Hay abundantes referencias en los escritos clásicos griegos a los conocimientos de los egipcios, a los que algunos griegos llegaron a considerar erróneamente como los fundadores de la ciencia, en particular de la agrimensura, la astronomía y la aritmética. Muchos griegos viajaron a Egipto para estudiar y conocer sus gentes, mientras otros visitaban Babilonia, y allí aprendieron su matemática y otras ciencias.

La influencia de Egipto y de Babilonia seguramente fue muy sensible en Mileto, una importante ciudad jónica en las costas de Asia Menor, en la que nacieron la filosofía, la matemática y las demás ciencias griegas. Mileto fue una importante y rica ciudad comercial del Mediterráneo, a cuyo puerto llegaban los barcos tanto de la Grecia continental como de Fenicia y Egipto; Babilonia estaba, en cambio, conectada a Mileto por medio de rutas de caravanas hacia el Este. Jonia cayó en manos de los persas hacia el 540 a. C., aunque Mileto conservó cierto grado de independencia. Una vez aplastado, el 494 a. C., el levantamiento jónico contra Persia, Jonia comenzó a perder su importancia. Volvió a formar parte de la Grecia propiamente dicha el 479 a. C., cuando los griegos derrotaron a los persas, pero para entonces la actividad cultural se había desplazado ya al territorio de la Grecia continental, con centro en Atenas.

A pesar de que la civilización griega antigua duró hasta el 600 d. C., aproximadamente, desde el punto de vista de la historia de la matemática conviene distinguir dos períodos: el clásico, que va desde el 600 al 300 a. C., y el alejandrino o helenístico, desde el 300 a. C. al 600 d. C. La adopción del alfabeto, que ya hemos mencionado, y el hecho de que el papiro estuviera disponible en Grecia durante el siglo VII a. C. quizás puedan explicar el florecimiento cultural que tuvo lugar hacia el 600 a. C. Indudablemente, el disponer de este material de escritura ayudó mucho a la difusión de las ideas.

2. Las fuentes generales

Sorprendentemente, las fuentes de las que procede nuestro conocimiento de la matemática griega son menos directas y fiables que las que tenemos de las matemáticas egipcia y babilónica, mucho más antiguas, debido a que no nos ha llegado ningún manuscrito original de los matemáticos griegos importantes de esa época. Una razón es, sin duda, la de que el papiro es un material de frágil consistencia; no obstante, los egipcios también utilizaron el papiro y, por suerte, se salvaron unos pocos de sus documentos matemáticos. Algunos de los voluminosos escritos griegos también podrían haber llegado hasta nosotros si no hubieran resultado destruidas sus grandes bibliotecas.

Nuestras fuentes principales para las obras matemáticas griegas son los códices bizantinos manuscritos en griego, escritos entre 500 y 1500 años después de que fueran escritas las obras griegas originales. Estos códices no suelen ser reproduciones literales, sino ediciones críticas, de manera que no podemos estar seguros de qué tipo de cambios hicieron los editores. También disponemos a veces de traducciones al árabe de las obras griegas, y de versiones latinas de estas traducciones al árabe; aquí, una vez más, no sabemos qué cambios pueden haber realizado los traductores ni hasta qué punto entendían correctamente los textos originales. Además, incluso los textos griegos utilizados por los autores árabes y bizantinos pudieron muy bien ser de autenticidad dudosa. Por ejemplo, aunque no disponemos del manuscrito de Herón, matemático griego de la época alejandrina, sí sabemos que hizo un cierto número de modificaciones en los *Elementos* de Euclides, dando demostraciones distintas y añadiendo nuevos casos de teoremas y sus recíprocos. Análogamente, Teón de Alejandría (finales del siglo IV d. C.) nos dice que modificó algunas secciones de los *Elementos* en su edición, y las versiones griegas y árabes que nos han llegado pueden provenir de tales versiones de los originales. Sin embargo, de una u otra forma, lo cierto es que disponemos de las obras de Euclides, de Apolonio, de Arquímedes, de Ptolomeo, de Diofanto y de otros matemáticos griegos. Muchos textos griegos escritos durante el período clásico y el alejandrino no han llegado hasta nosotros porque ya incluso en plena época griega se vieron superados por los escritos de estos autores.

Los griegos escribieron algunas historias de la matemática y de otras ciencias. Así, por ejemplo, Eudemo (siglo IV a. C.), miembro de la escuela aristotélica, escribió una historia de la aritmética, otra de la

geometría y otra de la astronomía, historias que, salvo fragmentos citados por escritores posteriores, se han perdido. La historia de la geometría trataba del período anterior a Euclides, y evidentemente sería inapreciable disponer de ella. Teofrasto (*c.* 372-*c.* 287 a. C.), otro discípulo de Aristóteles, escribió por su parte una historia de la física, que también se ha perdido, excepto unos cuantos fragmentos.

Además de los anteriores, tenemos dos importantes comentarios; Pappus (finales del siglo III d. C.) escribió su *Synagoge o Colección Matemática*, de la que conservamos casi su totalidad en una copia del siglo XII. Se trata de una exposición de la mayor parte de la obra de los matemáticos griegos clásicos y alejandrinos desde Euclides a Ptolomeo, complementada por un cierto número de lemas y teoremas que añade Pappus para facilitar su comprensión. Pappus mismo escribió también otra obra anterior titulada *Tesoro del Análisis*, que era una colección formada por las propias obras griegas. Esta obra se ha perdido, pero en el libro VII de su *Colección Matemática* nos resume lo que contenía el *Tesoro*.

El segundo comentarista importante es Proclo (410-485 d. C.), escritor muy prolífico. Proclo extrajo su material de los textos originales de los matemáticos griegos y de otros comentaristas anteriores. De las obras que nos han llegado, su *Comentario*, que estudia el libro I de los *Elementos* de Euclides, es la más importante. Según todos los indicios, Proclo trataba de escribir un comentario más extenso de los *Elementos*, pero al parecer nunca lo hizo. El *Comentario* contiene una de las tres citas atribuidas tradicionalmente a la historia de la geometría de Eudemo (véase la sección 10), pero probablemente tomadas de una modificación posterior. Este resumen concreto, el más largo de los tres, suele conocerse como el «sumario» de Eudemo. Proclo también nos dice algo sobre la obra de Pappus, de manera que, aparte de las ediciones y versiones posteriores de los clásicos griegos mismos, la *Colección Matemática* de Pappus y el *Comentario* de Proclo son las dos fuentes principales para la historia de la matemática griega.

Por lo que se refiere a las redacciones literales originales (aunque no, desde luego, los manuscritos), sólo disponemos de un fragmento relativo a la cuadratura de las lúnulas de Hipócrates, citado por Simplicio (primera mitad del siglo VI d. C.) y tomado de la *Historia de la Geometría* perdida de Eudemo, y un fragmento de Arquitas sobre la duplicación del cubo, y de los manuscritos originales nos han llegado algunos papiros escritos en la época alejandrina. Las fuentes

no estrictamente matemáticas, pero sí próximas, han resultado ser también de un enorme valor para la historia de la matemática griega. Por ejemplo, los filósofos griegos, especialmente Platón y Aristóteles, tenían mucho que decir sobre la matemática, y sus escritos han sobrevivido como las obras matemáticas mismas.

La reconstrucción de la historia de la matemática griega, basada en las fuentes que acabamos de mencionar, ha resultado una tarea gigantesca y complicada. A pesar de los grandes esfuerzos de los historiadores, todavía quedan lagunas en nuestros conocimientos y algunas conclusiones son discutibles; sin embargo, los hechos básicos están razonablemente claros.

3. Las escuelas principales del período clásico

Las contribuciones más importantes del período clásico son los *Elementos* de Euclides y las *Secciones Cónicas* de Apolonio. Para apreciar correctamente estas obras son necesarios algunos conocimientos de los grandes cambios experimentados en la naturaleza misma de la matemática y de los problemas con que se enfrentaron, y resolvieron, los griegos. Por otra parte, estas obras tan acabadas nos dan muy poca información sobre los trescientos años de actividad creadora que las precedieron o de las cuestiones que iban a ser vitales en la historia posterior.

La matemática clásica griega se desarrolló en diversos centros que se sucedían unos a otros, basándose cada uno en la obra de sus predecesores. En cada uno de estos centros, un grupo informal de matemáticos realizaba sus actividades dirigido por uno o más sabios. Este tipo de organización ha seguido funcionando en la época moderna, y su razón de ser se comprende fácilmente; hoy mismo, cuando un sabio importante se establece en un lugar concreto —normalmente una universidad—, otros estudiosos le siguen para aprender del maestro.

La primera de estas escuelas, la escuela jónica, fue fundada por Tales (*c.* 640-*c.* 546 a. C.) en Mileto. No sabemos con exactitud si Tales mismo enseñó a muchos otros, pero sí sabemos que los filósofos Anaximandro (*c.* 610-*c.* 547 a. C.) y Anaxímenes (*c.* 550-480 a. C.) fueron discípulos suyos. Anaxágoras (*c.* 500-*c.* 428 a. C.) perteneció también a esta escuela, y se supone que Pitágoras mismo (*c.* 585-*c.* 500 a. C.) pudo haber aprendido matemáticas de Tales; más tarde,

Pitágoras fundaría su propia e importante escuela al sur de Italia. Hacia finales del siglo VI, Jenófanes de Colofón, en Jonia, emigró a Sicilia y fundó a su vez un centro al que pertenecieron los filósofos Parménides (siglo V a. C.) y Zenón (siglo V a. C.). Estos últimos se establecieron en Elea, en el sur de Italia, ciudad a la que se trasladó la escuela, y por eso se conoció a este grupo como la escuela eleática. Los sofistas, que se mostraron activos desde mediados del siglo V en adelante, se concentraron principalmente en Atenas, ciudad en la que la escuela más famosa fue la Academia de Platón, de la que sería discípulo Aristóteles. La Academia tuvo una importancia sin precedentes para el pensamiento griego, sus discípulos y asociados fueron los más grandes filósofos, matemáticos y astrónomos de su época; y esta escuela conservaría su preeminencia en filosofía incluso después de que la capital de las matemáticas pasara a Alejandría. Eudoxo, que aprendió matemáticas principalmente de Arquitas de Tarento (Sicilia), fundó su propia escuela en Cízico, ciudad del norte de Asia Menor. Cuando Aristóteles abandonó la Academia de Platón, fundó a su vez otra escuela en Atenas, el Liceo; esta escuela ha recibido tradicionalmente el nombre de Escuela Peripatética. No todos los grandes matemáticos del período clásico pueden identificarse con una escuela concreta, pero para mayor claridad y coherencia estudiaremos la obra de cada matemático en relación con una escuela particular, incluso si su asociación a ella no fue demasiado estrecha.

4. La escuela jónica

El fundador de esta escuela y su figura más importante fue Tales. Aunque no sabemos nada con seguridad acerca de su vida y obra, Tales nació y vivió probablemente en Mileto; viajó mucho y durante algún tiempo vivió en Egipto, donde desarrolló actividades comerciales y, al parecer, aprendió mucho acerca de la matemática egipcia. Se supone, además, que fue un astuto comerciante que, aprovechando una buena cosecha de aceitunas, alquiló todas las almazaras de Mileto y Chíos para realquilarlas después a un precio más alto. Se dice que Tales anunció un eclipse de sol el año 585 a. C., pero esto es muy dudoso teniendo en cuenta los conocimientos astronómicos de la época.

Se le atribuye también el cálculo de las alturas de las pirámides comparando sus sombras con la de un bastón de altura conocida, en el

mismo instante, y mediante el mismo uso de los triángulos semejantes se supone que calculó la distancia de un buque a la playa. También se le ha atribuido la transformación de la matemática en una ciencia abstracta, y haber dado demostraciones deductivas de algunos teoremas, pero ambas cosas son de nuevo dudosas. Por último, se le ha atribuido a Tales el descubrimiento del poder de atracción de los imanes así como de la electricidad estática.

La escuela jónica sólo merece una breve mención por su contribución a la matemática propiamente dicha, pero su importancia para la filosofía, y la filosofía de la ciencia en particular, fue enorme (véase cap. 7, sec. 2). Esta escuela perdió su importancia a partir de la conquista de la región por los persas.

5. Los pitagóricos

La antorcha fue recogida por los pitagóricos que, habiendo aprendido de Tales, según se cuenta, fundaron sus propia escuela en Crotona, asentamiento griego en el sur de Italia. No conocemos ninguna obra escrita por los pitagóricos, y sólo sabemos de ellos por los escritos de otros, entre los que hay que incluir a Platón y Herodoto. Concretamente, apenas sabemos nada de la vida personal de Pitágoras y de sus seguidores, ni podemos tener la seguridad de qué hay que atribuirle a él personalmente o a sus discípulos. Por lo tanto, cuando se habla de la obra de los pitagóricos hay que tener en cuenta que en realidad nos estamos refiriendo a la obra del grupo entre el 585 a. C., presunta fecha de su nacimiento, y aproximadamente el 400 a. C. Filolao (siglo V a. C.) y Arquitas (428-347 a. C.) fueron dos miembros destacados de esta escuela.

Pitágoras nació en la isla de Samos, próxima a la costa de Asia Menor, y, después de algún tiempo estudiando con Tales en Mileto, viajó a otros países, entre ellos Egipto y Babilonia, donde asimiló su matemática al mismo tiempo que sus teorías místicas, y finalmente se estableció en Crotona. En esta ciudad fundó una especie de hermandad de tipo religioso, científico y filosófico. En realidad, era formalmente una escuela con un número limitado de miembros que aprendían de sus maestros. Las enseñanzas impartidas al grupo se mantenían en secreto por parte de los miembros, aunque, por lo que se refiere a la matemática y a la física, algunos historiadores niegan que existiera tal secreto. Se supone que los pitagóricos participaron en la política

de su ciudad aliándose con la facción aristocrática y terminaron siendo expulsados violentamente por el partido democrático o popular. Pitágoras huyó a la cercana Metaponto y allí murió, al parecer asesinado, hacia el 497 a. C. Sus seguidores se esparcieron por otras ciudades griegas y continuaron sus enseñanzas.

Una de las grandes contribuciones griegas al concepto mismo de la matemática fue el reconocimiento consciente y el énfasis puesto en el hecho de que los objetos matemáticos, números y figuras geométricas, son abstracciones, ideas producidas por la mente y claramente distintas de los objetos o imágenes físicas. Es cierto que incluso algunas civilizaciones primitivas, y con seguridad los egipcios y los babilonios, habían aprendido a pensar en los números separados de los objetos físicos, y, sin embargo, cabe preguntarse en qué medida eran conscientes del carácter abstracto de tal pensamiento. Por otra parte, los conceptos geométricos de todas las civilizaciones prehelénicas estaban decididamente ligados a la materia. Para los egipcios, por ejemplo, una recta no era más que una cuerda tensa o el borde de un terreno, y un rectángulo, su frontera.

El reconocimiento de que la matemática trabaja con abstracciones puede atribuirse con cierta seguridad a los pitagóricos. Sin embargo, puede que esto no fuera cierto desde el principio; Aristóteles nos dice, por ejemplo, que los pitagóricos consideraban a los números como los componentes últimos de los objetos materiales del mundo real [1]. Así pues, los números no tenían una existencia separada de los objetos sensibles. Cuando los primeros pitagóricos decían que todos los objetos estaban compuestos por números (enteros), o que los números eran la esencia del universo, lo entendían en sentido literal, porque los números eran para ellos como los átomos para nosotros. Se supone incluso que los pitagóricos de los siglos VI y V no distinguían realmente los números de los puntos geométricos, entendidos, naturalmente, como puntos extensos o esferas minúsculas. Sin embargo, Eudemo, según nos informa Proclo, decía que Pitágoras se remontó a principios más altos (que los de los egipcios y babilonios) y se ocupó de problemas abstractos de la inteligencia pura. Eudemo añade que Pitágoras fue el verdadero creador de la matemática pura, a la que convirtió en un arte liberal.

Los pitagóricos solían representar los números mediante puntos

[1] *Metafísica*, I, V, 986a y 986a 21, Ed. Gredos, Madrid, 1970.

en la arena o piedrecillas, clasificándolos según las formas de estas distribuciones de puntos o piedras. Así, los números 1, 3, 6, 10, etc. recibían el nombre de triangulares porque los puntos correspondientes podían distribuirse en forma de triángulo equilátero (fig. 3.1). El cuarto número triangular, el 10, ejerció una fascinación especial sobre los pitagóricos, siendo para ellos una especie de número sagrado, que tiene cuatro puntos en cada lado; el 4 era otro de sus números favoritos. Los pitagóricos comprobaron que las sumas 1, 1 + 2, 1 + 2 + 3, y así sucesivamente, daban lugar a los números triangulares y que $1 + 2 + ... + n = n \cdot (n + 1)/2$.

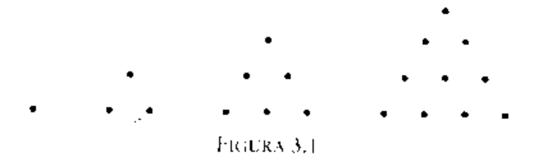


FIGURA 3.1

Los números 1, 4, 9, 16, ... recibieron el nombre de números cuadrados debido a que sus puntos pueden distribuirse formando cuadrados (fig. 3.2). Los números compuestos (o no primos) que no eran cuadrados perfectos recibían el nombre de oblongos.

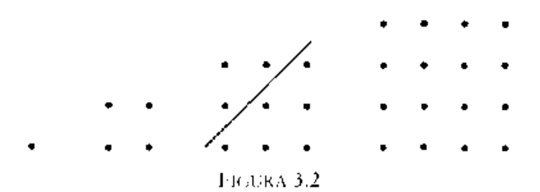


FIGURA 3.2

A partir de las distribuciones geométricas de los puntos aparecían como evidentes ciertas propiedades de los números enteros: por ejemplo, trazando la recta que aparece en la tercera ilustración de la figura 3.2 se descubre que la suma de dos números triangulares consecutivos es un número cuadrado. Esto es verdad en general, como podemos ver en notación moderna:

$$\frac{n(n + 1)}{2} + \frac{(n + 1)(n + 2)}{2} = (n + 1)^2.$$

Sin embargo, es dudoso que los pitagóricos pudieran demostrar esta conclusión general.

Para pasar de un número cuadrado al siguiente, los pitagóricos seguían el esquema que aparece en la figura 3.3; los puntos a la derecha y bajo las rectas en la figura forman lo que ellos llamaban un gnomon. Simbólicamente, lo que descubieron era que $n^2 + (2n + 1) = (n + 1)^2$. Además, si partimos del 1 y añadimos el gnomon 3 y después el gnomon 5, y así sucesivamente, lo que tenemos es, en nuestro simbolismo, $1 + 3 + 5 + ... + (2n - 1) = n^2$.

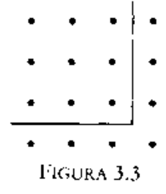


FIGURA 3.3

Con respecto a la palabra «gnomon», probablemente significó al principio, en Babilonia, una varilla vertical cuya sombra marcaba la hora. En la época de Pitágoras significaba una escuadra de carpintero, y esta es la forma del gnomon anterior. También significaba lo que queda de un cuadrado al cortar otro cuadrado más pequeño de una de sus esquinas, y más tarde, con Euclides, significó lo que queda de un paralelogramo al cortar otro más pequeño de una de sus esquinas, siempre que éste fuera semejante al primero (fig. 3.4).

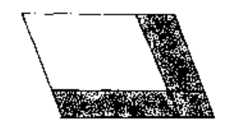


FIGURA 3.4. *El área sombreada es el gnomon.*

Los pitagóricos estudiaron también los números poligonales, tales como los pentagonales, hexagonales y otros. Como se ve en la figura 3.5, donde cada punto representa una unidad, el primer número pentagonal es el 1; el segundo, cuyos puntos forman los vértices de un pentágono, es el 5; el tercero es $1 + 4 + 7 = 12$, y así sucesivamente. El n-simo número pentagonal es, en nuestra notación, $(3n^2 - n)/2$. Análogamente, los números hexagonales (fig. 3.6) son 1, 6, 15, 28, ..., y en general $2n^2 - n$.

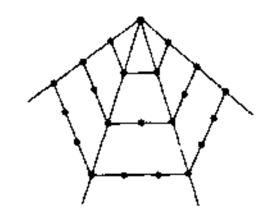


FIGURA 3.5. *Números pentagonales.*

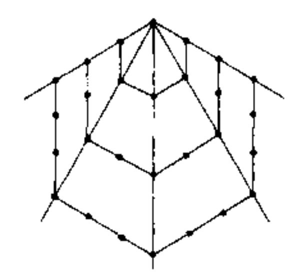


FIGURA 3.6. *Números hexagonales.*

Se llamó número perfecto a todo aquel que es igual a la suma de sus divisores, incluido el 1, pero no el propio número; por ejemplo, 6, 28, 496. A los que excedían a la suma de sus divisores se los llamó excesivos, y a los que eran menores que dicha suma, defectivos. A dos números se los llamó amigos cuando cada uno de ellos era igual a la suma de los divisores del otro, por ejemplo, 284 y 220.

Los pitagóricos descubrieron una regla para construir ternas de números enteros que pudieran ser lados de un triángulo rectángulo, sobre las cuales volveremos más adelante. Así, descubieron que si m es impar, entonces m, $(m^2 - 1)/2$ y $(m^2 + 1)/2$ constituyen una de estas ternas. Sin embargo, esta regla solamente da algunas de ellas. Cualquier terna de números enteros que represente los lados de un triángulo rectángulo recibe el nombre de terna pitagórica.

Los pitagóricos estudiaron los números primos, las progresiones, y ciertos tipos de razones y proporciones que encerraban para ellos una belleza especial. Así, si p y q son dos números, la media aritmética A es $(p + q)/2$, la media geométrica G es $\sqrt{pq}$, y la media armónica H, que es el recíproco de la media aritmética de $1/p$ y $1/q$, es $2pq/(p + q)$. Ahora bien, como puede comprobarse fácilmente, G es la media geométrica de A y H. La proporción $A/G = G/H$ recibió el nombre

de proporción perfecta, y la $p : (p + q)/2 = 2pq/(p + q) : q$ el de proporción musical.

Para los pitagóricos, los números eran únicamente los números enteros y una razón entre dos números no era una fracción y, por lo tanto, otro tipo de número como en la época moderna. Las fracciones concretas, utilizadas para expresar partes de una unidad monetaria o de una medida, se utilizaban evidentemente en el comercio, pero tales usos comerciales de la aritmética quedaban fuera del marco de la matemática griega propiamente dicha. Por lo tanto, los pitagóricos se vieron desagradablemente sorprendidos por el descubrimiento de que algunas razones, por ejemplo, la razón de la hipotenusa de un triángulo rectángulo isósceles a un cateto o, lo que es lo mismo, de la diagonal al lado de un cuadrado, no podían expresarse por medio de números enteros. Dado que los pitagóricos se habían dedicado a estudiar las ternas de números enteros que podían ser lados de un triángulo rectángulo, lo más probable es que descubrieran estas nuevas razones en el mismo contexto. Llamaron razones conmensurables a las que se podían expresar por medio de números enteros, lo que significaba que las dos cantidades venían medidas por una unidad común, y a las que no eran expresables de esa manera, razones inconmensurables; por lo tanto, lo que nosotros expresamos de la forma $\sqrt{2}/2$ es una razón inconmensurable. Una razón entre magnitudes inconmensurables recibió el nombre de *αλογος* (*alogos* o inexpresable), aunque también se utilizó el término *αρρητος* (*arretos* o que no tiene razón). El descubrimiento de las razones inconmensurables se atribuye a Hipaso de Metaponto (siglo V a. C.), suponiéndose que los pitagóricos se encontraban navegando en el mar en esa época y que lanzaron a Hipaso por la borda como castigo por haber introducido en el universo un elemento que negaba la teoría pitagórica de que todos los fenómenos del universo se podían reducir a números enteros y sus razones.

La demostración dada por los pitagóricos de la inconmensurabilidad de $\sqrt{2}$ con 1 procedía, según Aristóteles, por *reductio ad absurdum*, es decir, por el método de demostración indirecta. Concretamente, la demostración mostraba que si la hipotenusa fuera conmensurable con el cateto, entonces el mismo número tendría que ser a la vez par e impar, y transcurre de la manera siguiente: sea la razón de la hipotenusa al cateto en un triángulo rectángulo isósceles α/β y consideremos expresada esta razón mediante los números más pequeños posibles. Entonces $\alpha^2 = 2\beta^2$ por el teorema de Pitágoras, y dado

que α^2 es par, α debe serlo también, puesto que el cuadrado de cualquier número impar es impar [2]. Ahora bien, la razón α/β estaba expresada en sus términos mínimos, luego β tiene que ser impar; como α es par, sea $\alpha = 2\gamma$, luego $\alpha^2 = 4\gamma^2 = 2\beta^2$, luego $\beta^2 = 2\gamma^2$, es decir, β^2 par y β también par. Pero β era impar por lo anterior y hemos llegado a una contradicción.

Esta demostración, que es la misma, desde luego, que la demostración moderna de la irracionalidad de $\sqrt{2}$, aparecía incluida en antiguas ediciones de los *Elementos* de Euclides como proposición 117 del libro X. Sin embargo lo más probable es que no apareciera en el texto original de Euclides y, por lo tanto, se suele omitir en las ediciones modernas.

En la matemática moderna las razones inconmensurables se expresan por medio de números irracionales, pero los pitagóricos nunca habrían aceptado tales números. Los babilonios trabajaron, de hecho, con tales números mediante aproximaciones, aunque probablemente no sabían que tales aproximaciones sexagesimales fraccionarias nunca podían ser exactas, así como tampoco los egipcios llegaron a reconocer el carácter distinto de los irracionales. Los pitagóricos, al menos, reconocieron que las razones inconmensurables son de un tipo completamente diferente de las conmensurables.

Este descubrimiento planteó un problema central en la matemática griega. Hasta este momento los pitagóricos habían identificado número y geometría, pero la existencia de razones inconmensurables destruía esta identificación. No cesaron de considerar todo tipo de longitudes, áreas y razones en geometría, pero se restringieron a considerar razones numéricas únicamente para el caso conmensurable. La teoría de proporciones para razones inconmensurables y para todo tipo de magnitudes se debe a Eudoxo, cuya obra estudiaremos más adelante.

Hay algunos otros resultados geométricos descubiertos también por los pitagóricos. El más famoso es, desde luego, el mismísimo teorema de Pitágoras, un teorema clave para la geometría euclídea, pero también se le atribuyen muchos de los teoremas que conocemos sobre triángulos, rectas paralelas, polígonos, círculos, esferas y los poliedros regulares. Concretamente, sabían que la suma de los ángu-

---

[2] Todo número impar puede expresarse de la forma $2n + 1$, para algún n. Entonces se tiene que $(2n + 1)^2 = 4n^2 + 4n + 1$, que siempre es impar de nuevo.

los de un triángulo es de 180 grados, y entre otros resultados conocían una teoría restringida de figuras semejantes y el hecho de que un plano puede ser recubierto por triángulos equiláteros, cuadrados y hexágonos regulares.

Los pitagóricos empezaron a estudiar un tipo de problemas conocidos con el nombre de aplicación de áreas. El más sencillo de ellos era el de construir un polígono de área igual a uno dado y semejante a otro dado. Otro consistía en construir una figura concreta con un área que excedía o resultaba defectuosa de otra en un área dada. La forma más importante del problema de aplicación de áreas es: dado un segmento, construir sobre una parte de él o sobre él mismo extendido un paralelogramo igual en área a una figura rectilínea dada y resultando deficiente (en el primer caso) o excediendo (en el segundo caso) en un paralelogramo semejante a uno dado. Más adelante discutiremos en detalle estas aplicaciones de áreas al estudiar la obra de Euclides.

La contribución más esencial de los griegos a la matemática fue su insistencia en que todos los resultados matemáticos deberían ser establecidos deductivamente a partir de un sistema explícito de axiomas. Por lo tanto, se plantea la cuestión de si los pitagóricos demostraban ya sus resultados geométricos. No podemos dar una respuesta definitiva, pero es muy dudoso que los pitagóricos del período antiguo o medio exigieran demostraciones deductivas, explícitas o implícitas, basadas en un sistema de axiomas de cualquier tipo. Proclo nos asegura que demostraron el teorema de la suma de los ángulos de un triángulo, pero esto puede ser debido a los pitagóricos tardíos. La cuestión acerca de si demostraron el teorema de Pitágoras ha sido muy discutida, y la conclusión generalmente aceptada es la de que probablemente no. Es relativamente fácil demostrarlo utilizando resultados sobre triángulos semejantes, pero lo cierto es que los pitagóricos no tenían una teoría completa de la semejanza. La demostración dada en la proposición 47 del libro I de los *Elementos* de Euclides (cap. 4, sec. 4) es difícil porque no utiliza la teoría de figuras semejantes, y se trata de una demostración que Proclo atribuye a Euclides mismo. La conclusión más verosímil acerca de la presencia de demostraciones en la geometría pitagórica es la de que durante la mayor parte de la vida de la escuela los miembros justificaban sus resultados sobre la base de casos especiales, análogamente a como hacían en aritmética. Sin embargo, en la época de los pitagóricos tardíos, es decir, hacia el 400 a. C., el *status* de la demostración había cambiado ya debido a otros

desarrollos; así pues, estos miembros tardíos de la hermandad pudieron haber dado ya demostraciones rigurosas.

6. La escuela eleática

El descubrimiento pitagórico de las razones inconmensurables introdujo en escena una dificultad que preocupó a los griegos, a saber, la relación entre lo discreto y lo continuo. Los números enteros representan objetos discretos y una razón conmensurable representa una relación entre dos colecciones de objetos discretos o entre dos longitudes que admiten una unidad de medida común, de manera que cada una de ellas es una colección discreta de unidades. Sin embargo, las longitudes en general no son colecciones discretas de unidades, y este es el motivo de que aparezcan las razones de longitudes inconmensurables. En otras palabras, longitudes, áreas, volúmenes, tiempos y otras cantidades son continuas. Nosotros diríamos que los segmentos rectilíneos, por ejemplo, pueden tener longitudes racionales o irracionales en términos de alguna unidad concreta, pero los griegos no dieron este paso.

El problema de la relación entre lo discreto y lo continuo fue puesto en evidencia por Zenón, que vivió en la ciudad de Elea, al sur de Italia. Zenón nació entre el año 495 y el 480 a. C., y era más bien un filósofo que un matemático, del que, al igual que de su maestro Parménides, se dice que fue inicialmente un pitagórico. Zenón propuso un cierto número de paradojas, cuatro de las cuales tratan del movimiento, cuyo objeto no está del todo claro debido a nuestro conocimiento incompleto de la historia de la filosofía griega. Se dice que con ellas pretendía defender a Parménides, que había sostenido que el movimeinto o el cambio en general es imposible, y también que trataba de atacar a los pitagóricos, que creían en unidades extensas pero indivisibles, los puntos de la geometría. No sabemos exactamente lo que dijo Zenón, sino que nos vemos obligados a apoyarnos en citas de Aristóteles, que menciona a Zenón con objeto de criticarlo, y de Simplicio, que vivió en el siglo VI d. C. y que basaba sus afirmaciones en los escritos de Aristóteles.

Las cuatro paradojas sobre el movimiento son distintas, pero el argumento importante probablemente consistía en las cuatro consideradas en bloque. En la época en que vivió Zenón había dos concepciones opuestas del espacio y del tiempo: una, que el espacio y el tiempo

son indefinidamente divisibles, en cuyo caso el movimiento resultaría continuo y «liso»; y la otra, que el espacio y el tiempo están formados por pequeños intervalos indivisibles (como en el cine), en cuyo caso el movimiento consistiría en una sucesión de minúsculos saltos espasmódicos. Los argumentos de Zenón están dirigidos contra ambas teorías, las dos primeras paradojas contra la primera, y las otras dos contra la segunda. La primera paradoja de cada pareja considera el movimiento de un único cuerpo, y la segunda el movimiento relativo de un cuerpo con respecto a otro.

Aristóteles formula en su *Física* la primera paradoja, llamada de Dicotomía, de la manera siguiente: «La primera afirma la no existencia del movimiento basándose en que lo que está en movimiento debe alcanzar la posición a medio camino antes de alcanzar su meta.» Esto significa que para atravesar *AB* (fig. 3.7) hay que alcanzar primero la posición *C*; para llegar a *C* hay que llegar primero a *D*, y así sucesivamente. En otras palabras, sobre la hipótesis de que el espacio es indefinidamente divisible y por lo tanto que una longitud finita contiene un número infinito de puntos, es imposible cubrir incluso una longitud finita en un tiempo finito.

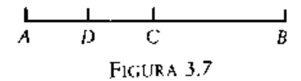


FIGURA 3.7

Aristóteles, intentando refutar a Zenón, dice que hay dos sentidos en los que una cosa puede ser infinita: en extensión o en divisibilidad. En un tiempo finito se puede establecer contacto con infinitas cosas en el sentido de la divisibilidad, ya que en este sentido el tiempo también es infinito; y así una extensión finita de tiempo puede ser suficiente para cubrir una longitud finita. Según otros, este argumento de Zenón ha sido construido para poner de relieve que al atravesar una longitud finita hay que recorrer un número infinito de puntos y así alcanzar el final de algo que esencialmente no tiene final.

La segunda paradoja lleva el nombre de Aquiles y la Tortuga. Según Aristóteles: «Afirma que el objeto que se mueve más lentamente no puede ser alcanzado por el más rápido ya que el perseguidor debe llegar primero al punto del cual partió el perseguido, de manera que el más lento necesariamente estará siempre en cabeza. El argumento es análogo al de la Dicotomía, pero la diferencia radica en que no dividimos en mitades las distancias que se han de recorrer.» Aristóteles dice entonces que si el objeto que se mueve lentamente

cubre una distancia finita, puede ser superado por la misma razón que daba al responder a la primera paradoja.

Las otras dos paradojas están dirigidas contra el movimiento «cinematográfico». La tercera, llamada de la Flecha, nos la presenta Aristóteles como sigue: «La tercera paradoja que formuló Zenón es la de que una flecha moviéndose está en reposo; él llega a esta conclusión a partir de la hipótesis de que el tiempo está constituido por instantes. Si no fuera por esta hipótesis no habría tal conclusión.» Según Aristóteles, lo que dice Zenón es que en cualquier instante durante su movimiento la flecha ocupa una posición determinada y por lo tanto está en reposo. Así pues, no puede estar en movimiento. Aristóteles afirma que esta paradoja falla si no admitimos las unidades de tiempo indivisibles.

La cuarta paradoja, llamada del Estadio o de las Filas en Movimiento, la fórmula Aristóteles con estas palabras: «La cuarta consiste en el argumento acerca de un conjunto de cuerpos moviéndose en una carrera y cruzándose con otro conjunto de cuerpos en número igual y moviéndose en dirección opuesta, el primero partiendo del final y el otro de punto medio y moviéndose ambos con igual velocidad; Zenón concluye que de esto se sigue que la mitad del tiempo es igual a su doble. El error consiste en suponer que dos cuerpos moviéndose a velocidades iguales consumen tiempos iguales en cruzarse, el primero con un cuerpo que está en movimiento y el segundo con otro de igual tamaño que está en reposo, hipótesis que es falsa.»

La interpretación más probable de la cuarta paradoja de Zenón podría formularse de la manera siguiente: supongamos que tenemos tres filas de soldados *A*, *B* y *C* (fig. 3.8), y que en la mínima unidad de tiempo toda la fila *B* se mueve una posición hacia la izquierda, mientras que en el mismo tiempo la fila *C* se mueve una posición hacia la derecha. Entonces, relativamente a *B*, *C* se ha movido dos posiciones, y por lo tanto ha debido haber una unidad de tiempo menor al cabo de la cual *C* estaría una posición a la derecha de *B*, o bien la mitad de la unidad de tiempo resultaría ser igual a la unidad misma.

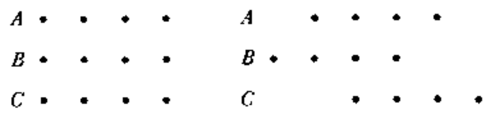


FIGURA 3.8

Es posible que Zenón intentara simplemente señalar que la velocidad es relativa. La velocidad de *C* relativa a *B* no es la misma que la relativa a *A*. O bien puede haber querido indicar que no hay un espacio absoluto al que referir las velocidades. Aristóteles dice que la falacia de Zenón consiste en suponer que las cosas que se mueven con la misma velocidad emplean el mismo tiempo en adelantar a un objeto en movimiento y a un objeto fijo. Ni el argumento de Zenón ni la respuesta de Aristóteles son claros, pero si suponemos que la paradoja consiste en un ataque a los intervalos mínimos indivisibles y a los segmentos mínimos indivisibles de espacio, que es lo que Zenón intentaba atacar, entonces su argumentación tiene perfecto sentido.

Podemos incluir a Demócrito (*c*. 460-*c*. 370 a. C.) de Abdera, en Tracia, entre los eleáticos. Es fama que Demócrito fue un hombre de gran sabiduría que trabajó en muy diversos campos, incluida la astronomía. Dado que perteneció a la escuela de Leucipo, y éste fue un discípulo de Zenón, muchas de las cuestiones matemáticas que estudió Demócrito debieron venir sugeridas por ideas de Zenón. Escribió obras de geometría, de aritmética y de líneas y sólidos continuos; concretamente, las obras geométricas pudieron muy bien haber estado entre los antecedentes de los *Elementos* de Euclides.

Arquímedes nos dice que fue Demócrito quien descubrió que los volúmenes de un cono y de una pirámide son iguales a 1/3 de los volúmenes del cilindro y prisma que tienen la misma base y la misma altura, pero que las demostraciones de estos dos resultados se deben a Eudoxo. Demócrito consideraba al cono como una serie de capas muy finas e indivisibles (fig. 3.9), pero se encontró enfrentado con la dificultad de que si las capas fueran todas iguales darían un cilindro, mientras que si fueran distintas la superficie del cono no sería lisa.

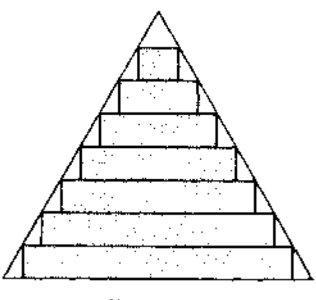


FIGURA 3.9

7. Los sofistas

Después de la derrota final de los persas en Micala el 479 a. C., Atenas se convirtió en la ciudad más importante de una liga de ciudades griegas, y en un floreciente centro comercial. La riqueza acumulada en el comercio, que hizo de Atenas la ciudad más rica de su época, fue utilizada por el famoso gobernante Pericles para reconstruir y adornar la ciudad. Jónicos, pitagóricos y todo tipo de intelectuales se vieron atraídos a Atenas, donde se ponía un especial énfasis en el razonamiento abstracto con el fin de extender el dominio de la razón tanto a la naturaleza como al hombre mismo.

La primera escuela ateniense, llamada la de los sofistas, incluía eruditos maestros en gramática, retórica, dialéctica, elocuencia, moral y —lo que más nos interesa a nosotros— geometría, astronomía y filosofía. Uno de sus objetivos principales era el de usar la matemática para entender el funcionamiento del universo.

Muchos de los resultados matemáticos obtenidos fueron subproductos de los intentos de resolver los tres famosos problemas de construcciones: construir un cuadrado de área igual a un círculo dado; construir la arista de un cubo de volumen doble que otro de arista dada; y trisecar un ángulo cualquiera: todo ello debía ser realizado con regla y compás únicamente.

Se han dado diversas explicaciones sobre el origen de estos famosos problemas de construcciones. Por ejemplo, una versión del origen del problema de la duplicación del cubo, encontrada en una obra de Eratóstenes (*c.* 284-192 a. C.), nos cuenta que los habitantes de Delos, bajo el azote de una peste, consultaron al oráculo sobre la manera de librarse de ella, a lo que el oráculo respondió que debían construir un altar de tamaño doble del que ya existía, de forma cúbica. Los habitantes de Delos comprobaron que duplicando la arista no se duplicaba el volumen, y se dirigieron a Platón, quien les dijo que el dios del oráculo no había contestado así porque quisiera o necesitara un altar doble, sino para censurar a los griegos por su indiferencia con respecto a la matemática y su falta de respeto por la geometría. Plutarco también nos cuenta la misma historia.

En realidad, estos problemas de construcciones eran generalizaciones de otros problemas ya resueltos por los griegos. Dado que cualquier ángulo podía ser bisecado, era natural plantearse la trisección. Y dado que la diagonal de un cuadrado es el lado de un cuadrado de área doble que el original, el problema correspondiente para el

cubo resulta también muy natural. El problema de cuadrar el círculo es un caso típico de muchos problemas griegos de construir una figura de forma dada y de área igual a otra figura dada. Otro problema no tan famoso fue el de la construcción de los polígonos regulares de 7 o más lados; aquí, de nuevo, la construcción del cuadrado, del pentágono y del hexágono regulares sugirieron la etapa siguiente.

Se han dado diversas explicaciones acerca de la restricción a la regla y el compás como instrumentos. La línea recta y la circunferencia eran, a los ojos de los griegos, las figuras básicas, traducidas físicamente en la regla y el compás, y por lo tanto se consideraron preferibles las construcciones con estos dos instrumentos. También se ha esgrimido la razón de que Platón puso objeciones a otros instrumentos mecánicos porque hacían intervenir demasiado el mundo de los sentidos en lugar del mundo de las ideas, que él consideraba como primario. Es muy probable, sin embargo, que en el siglo V la restricción a la regla y el compás no fuera tan rígida, pero, como veremos, las construcciones jugaron un papel vital en la geometría griega, y los axiomas de Euclides las limitaron a las que se pueden hacer con regla y compás; por lo tanto, desde ese momento en adelante tal restricción puede haberse tomado con más seriedad. Pappus, por ejemplo, nos dice que si una construcción puede hacerse con regla y compás, cualquier otra solución utilizando medios distintos no es satisfactoria.

El primer intento conocido de resolver uno de los tres famosos problemas se debió al jonio Anaxágoras, quien se supone trató de resolver la cuadratura del círculo mientras se encontraba en prisión; no sabemos nada más sobre el caso. Otro de los intentos más famosos fue el de Hipias de Elis, una ciudad del Peloponeso. Hipias fue uno de los sofistas más importantes, nacido hacia el 460 a. C. y contemporáneo de Sócrates.

Intentando trisecar el ángulo inventó Hipias una nueva curva, que desgraciadamente no es constructible con regla y compás. Esta curva se llama la cuadratriz o trisectriz, y se genera de la manera siguiente: sea AB (fig. 3.10) un segmento d que gira en el sentido de las agujas de un reloj alrededor de A a una velocidad constante, hasta ocupar la posición AD. Durante el mismo tiempo BC se mueve hacia abajo manteniéndose paralela a sí misma y a una velocidad constante hasta alcanzar la posición AD. Supongamos que AB se encuentra en la posición AD' al mismo tiempo que BC ocupa la posición $B'C'$, y sea E' el punto de interseción de AD' con $B'C'$. Entonces E' es un punto

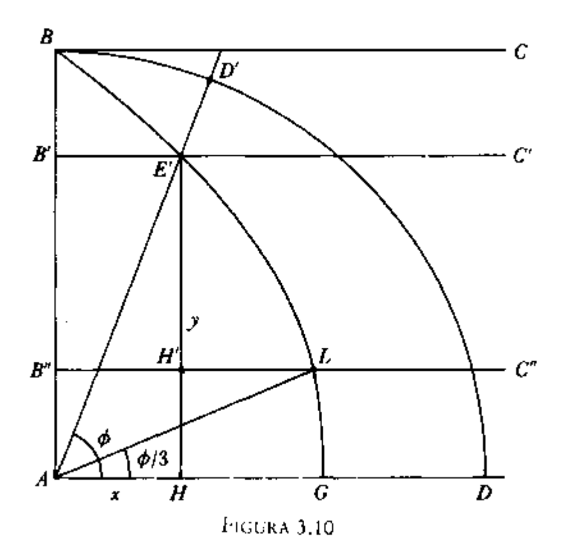


FIGURA 3.10

genérico de la cuadratriz $BE'G$. El punto límite G es el final de la cuadratriz [3].

La ecuación de la cuadratriz en coordenadas cartesianas rectangulares puede obtenerse de la manera siguiente: supongamos que AD' alcanza AD en alguna fracción t/T del tiempo total T que invierte AB en alcanzar AD. Como AD' y $B'C'$ se mueven con velocidades constantes, $B'C'$ recorre la parte $E'H$ de BA en la misma fracción del tiempo total; por tanto,

$$\frac{\phi}{\pi/2} = \frac{E'H}{BA}$$

Si representamos $E'H$ por y y BA por a entonces

$$\frac{\phi}{\pi/2} = \frac{y}{a} \qquad (1)$$

[3] El punto G no puede obtenerse directamente a partir de la definición de la curva, porque AB alcanza la posición AD al mismo tiempo que BC, y por lo tanto no está determinado el punto de intersección de las dos rectas. G sólo puede obtenerse como límite de puntos anteriores de la cuadratriz. Utilizando el cálculo infinitesimal puede demostrarse que $AG = 2a/\pi$, donde $a = AB$.

o bien

$$y = a\,\phi\,\frac{2}{\pi}$$

Pero si $AH = x$, entonces

$$\phi = \text{arc tg}\,\frac{y}{x}\,,$$

por lo tanto

$$y = \frac{2a}{\pi}\,\text{arc tg}\,\frac{y}{x}$$

o bien

$$y = x\,\text{tg}\,\frac{\pi y}{2a}\,.$$

La curva, si fuera constructible, podría ser utilizada para trisecar cualquier ángulo agudo. En efecto, sea ϕ tal ángulo; entonces dividamos y en tres partes iguales de manera que $E'H' = 2H'H$. Tracemos $B''C''$ por H' y supongamos que corta a la cuadratriz en L. Si trazamos AL, entonces $\sphericalangle LAD = \phi/3$, puesto que por el razonamiento que nos condujo a (1),

$$\frac{\sphericalangle LAD}{\pi/2} = \frac{H'H}{a}$$

o bien

$$\frac{\sphericalangle LAD}{\pi/2} = \frac{y/3}{a}\,.$$

Pero por (1)

$$\frac{\phi}{\pi/2} = \frac{y}{a}$$

luego

$$\sphericalangle LAD = \frac{\phi}{3}.$$

Otro descubrimiento famoso que se obtuvo del estudio de los problemas de construcciones fue el que hizo Hipócrates de Chíos (siglo V a. C.), el más famoso matemático de este siglo, al que no hay que confundir con su contemporáneo Hipócrates de Cos, padre de la medicina griega. Hipócrates floreció en Atenas durante la segunda mitad del siglo; no se trataba de un sofista, sino más bien de un pitagórico. Se le atribuye la idea de ordenar los teoremas de manera que los posteriores se puedan demostrar a partir de los anteriores, de una manera familiar para nosotros desde Euclides. También se le atribuye la introducción en matemáticas del método de demostración indirecto. Al parecer, escribió un texto de geometría titulado *Elementos,* que se ha perdido.

Hipócrates no resolvió el problema de la cuadratura del círculo, evidentemente, pero sí resolvió otros relacionados con él. Sea, por ejemplo, ABC un triángulo rectángulo isósceles (fig. 3.11) inscrito en la semicircunferencia de centro O. Sea AEB la semicircunferencia de diámetro AB. Entonces

$$\frac{\text{Area del semicírculo } ABC}{\text{Area del semicírculo } AEB} = \frac{AC^2}{AB^2} = \frac{2}{1}.$$

Por lo tanto, el área $OADB$ será igual al área del semicírculo AEB; si restamos a ambos el área común ADB entonces el área de la lúnula (o

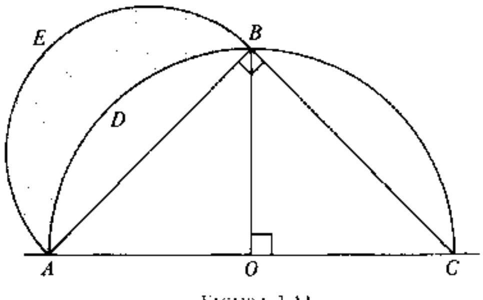


FIGURA 3.11

región sombreada) será igual al área del triángulo *AOB*. Así pues, el área de la lúnula, que es una figura limitada por arcos, es igual al área de una figura rectilínea; dicho con otras palabras, una figura curvilínea ha quedado reducida a otra rectilínea. Este resultado es una cuadratura, es decir, se ha calculado de manera efectiva un área curvilínea porque es igual a un área limitada por líneas rectas, y ésta puede ser calculada.

En su demostración hace uso Hipócrates del hecho de que dos círculos son entre sí como los cuadrados construidos sobre sus diámetros. Es muy dudoso que Hipócrates pudiera dar realmente una demostración de este hecho, puesto que tal demostración depende del método de exhausción inventado más tarde por Eudoxo.

Hipócrates consiguió cuadrar otras tres lúnulas, trabajo que conocemos a través de los escritos de Simplicio, y se trata del único fragmento de la matemática *clásica* griega que nos ha llegado en su redacción original.

También demostró Hipócrates que el problema de la duplicación del cubo puede reducirse a encontrar dos medias proporcionales entre la arista dada y su doble. En nuestra notación algebraica, sean x e y tales que

$$\frac{a}{x} = \frac{x}{y} = \frac{y}{2a}.$$

Entonces

$$x^2 = ay \quad \text{y} \quad y^2 = 2ax.$$

Y como $y = x^2/a$, de la segunda ecuación se obtiene que $x^3 = 2a^3$, que es la respuesta deseada, y que no puede construirse con regla y compás. Desde luego, Hipócrates debió razonar geométricamente, de una manera que veremos más clara cuando estudiemos las *Secciones Cónicas* de Apolonio.

Otra idea muy importante fue la que se les ocurrió a los sofistas Antifón (siglo V a. C.) y Brisson (*c.* 450 a. C.). Al intentar cuadrar el círculo se le ocurrió a Antiphon la idea de aproximarse a dicha figura por medio de polígonos inscritos de número de lados cada vez mayor. Y Brisson incorporó la idea de utilizar polígonos circunscritos. Antifón, por su parte, vino a sugerir además que el círculo podría ser considerado como un polígono de un número infinito de lados; más

adelante veremos cómo utilizó Eudoxo estas ideas en su método de exhausción (cap. 4 sec. 9).

8. La escuela platónica

La escuela platónica sucedió a los sofistas a la cabeza de la actividad matemática. Sus precursores inmediatos, Teodoro de Cirene, en el norte de Africa (nacido hacia el 470 a. C.), y Arquitas de Tarento, en el sur de Italia (428-347 a. C.) fueron pitagóricos y maestros ambos de Platón, de manera que sus enseñanzas pudieron haber sido las que dieron lugar a la fuerte influencia pitagórica en toda la escuela de Platón.

A Teodoro se atribuye el haber demostrado que las razones que nosotros representamos por $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$, ..., $\sqrt{17}$ son todas inconmensurables con la unidad. Arquitas, por su parte, introdujo la idea de considerar una curva como generada por un punto en movimiento, y una superficie generada por una curva en movimiento. Usando esta idea resolvió el problema de la duplicación del cubo hallando dos medias proporcionales entre dos cantidades dadas; estas medias proporcionales se construyen geométricamente hallando la intersección de tres superficies: la que genera una circunferencia girando alrededor de una tangente, un cono y un cilindro. La construcción es bastante complicada, y no entraremos aquí en los detalles. Arquitas escribió también sobre mecánica matemática, diseñó máquinas, estudió el sonido y contribuyó a las escalas musicales mediante ciertos inventos y algo de teoría.

La escuela platónica estuvo encabezada, naturalmente, por Platón, e incluyó entre sus miembros a Menecmo y su hermano Dinostrato (siglo IV a. C.) y a Teeteto (*c.* 415-*c.* 369 a. C.). A muchos otros miembros los conocemos sólo de nombre.

Platón (427-347 a. C.) nació en una familia distinguida, y de joven tuvo ambiciones políticas, pero la suerte de Sócrates le convenció de que no había lugar en la política para un hombre de conciencia. Viajó a Egipto y visitó a los pitagóricos en el sur de Italia; la influencia pitagórica pudo producirse a través de estos contactos. Hacia el 387 a. C. fundó Platón su Academia en Atenas, la cual se parecía en muchos sentidos a una universidad moderna. La Academia disponía de terrenos, edificios, estudiantes, y allí daban cursos formalmente Platón y sus ayudantes. Durante el período clásico se vio especial-

mente favorecido el estudio de la filosofía y de la matemática, y aunque el principal centro matemático se desplazó a Alejandría hacia el 300 a. C., la Academia siguió manteniendo su preeminencia en filosofía durante todo el período alejandrino. En total duró casi 900 años hasta su cierre por orden del emperador cristiano Justiniano el año 529 d. C., porque enseñaba «conocimientos paganos y perversos».

Platón, que fue uno de los hombres más sabios de su época, no era matemático, pero su entusiasmo por la materia y la creencia en su importancia para la filosofía y para el entendimiento del universo hizo que animara a los matemáticos a cultivarla. Es notable, y así hay que destacarlo, que casi todas las obras matemáticas importantes del siglo IV se deban a amigos o discípulos de Platón. Platón mismo parece haber estado más interesado en mejorar y perfeccionar lo que ya se conocía.

Aunque no podemos estar seguros de en qué medida los conceptos de la matemática fueron considerados como abstracciones antes de la época de Platón, no cabe duda de que Platón y sus sucesores los consideraron así. Platón dice que los números y conceptos geométricos no tienen en sí nada material y son distintos de los objetos físicos. Así pues, los conceptos de la matemática son independientes de la experiencia y tienen una realidad propia; se los descubre, no se los inventa o crea, y esta distinción entre abstracciones y objetos materiales pudo tener su origen en Sócrates.

Una cita de la *República* de Platón puede servir para ilustrar la concepción contemporánea de los objetos matemáticos. Sócrates se dirige a Glaucón:

> Y toda la aritmética y el cálculo tienen que ver con el número.
>
> Si...
>
> Entonces este es un conocimiento del tipo que estamos buscando, que tiene un doble uso, militar y filosófico; pues el hombre de guerra debe aprender el arte de los números o no sabrá cómo disponer sus tropas, y el filósofo también, porque tiene que salir del mar del cambio y buscar el verdadero ser, y por lo tanto debe de ser un aritmético... Por lo tanto este es un tipo de conocimiento que la legislación puede prescribir adecuadamente, y debemos intentar persuadir a los que estén destinados a ser hombres principales de nuestro Estado para que aprendan aritmética, pero no sólo como aficionados, sino que deben proseguir ese estudio hasta ver la naturaleza de los números sólo con la mente; y no, una vez más, como los mercaderes o los tenderos al por menor, con la vista puesta en comprar o vender, sino por su

utilidad militar y para el alma misma, debido a que éste será el camino más fácil para ella de pasar del cambio a la verdad y el ser... Entiendo, como estaba diciendo, que la aritmética tiene un gran efecto de elevación, impulsando al alma a razonar sobre el número abstracto, y rechazando la introducción de objetos visibles o tangibles en el razonamiento... [4]

En otro contexto [5], se discuten los conceptos de la geometría. Hablando acerca de los matemáticos dice Platón: «Y no sabéis también que aunque hacen uso de las formas visibles y razonan acerca de ellas, no piensan en éstas, sino en los ideales a que ellas semejan... Pero están intentando realmente contemplar las cosas mismas, que sólo pueden ser vistas con los ojos de la mente».

Estas citas dejan claro que Platón y otros griegos para los que él habla valoraban las ideas abstractas y preferían las ideas matemáticas como preparación para la filosofía. Las ideas abstractas de las que se ocupa la matemática son afines a otras, tales como la bondad y la justicia, cuyo entendimiento es la meta de la filosofía de Platón. Así pues, la matemática es la preparación para el conocimiento del universo ideal.

¿Por qué preferían y subrayaban los griegos los conceptos abstractos de la matemática? No podemos contestar a esta pregunta, pero hay que observar que los primeros matemáticos griegos fueron filósofos y que los filósofos en general ejercieron una influencia formativa importante en el desarrollo de la matemática griega. Los filósofos están interesados en las ideas y muestran una propensión típica por las abstracciones en muchos campos; así, los filósofos griegos pensaron sobre la verdad, el bien, la caridad y la inteligencia, especularon sobre la sociedad ideal y el estado perfecto, y tanto los pitagóricos tardíos como los platónicos distinguieron claramente entre el mundo de las ideas y el de las cosas. Las relaciones en el mundo material estaban sujetas a cambios y no representaban por ello la verdad última, pero las relaciones en el mundo ideal eran inmutables y por lo tanto verdades absolutas; éstas eran pues el objeto propio del filósofo.

Platón, concretamente, creía que las idealizaciones perfectas de los objetos físicos son la auténtica realidad; el mundo de las ideas y relaciones entre ellas es permanente, sin edad, incorruptible y universal, mientras que el mundo físico es una realización imperfecta del

---

[4] Libro VII, 525; Platón, *Diálogos*, vol. IV, Ed. Gredos, Madrid.

[5] *República*, libro VI, 510.

mundo ideal y está sujeto a la degradación. Por lo tanto, sólo el mundo ideal merece estudio y sólo se puede obtener un conocimiento infalible de las puras formas inteligibles. Sobre el mundo físico sólo podemos tener opiniones, y la ciencia física está condenada a verse hundida en el fango de un mundo de sensaciones.

No sabemos si los platónicos contribuyeron decisivamente a la estructura deductiva de la matemática, aunque sí sabemos que se interesaron por la demostración y la metodología del razonamiento. Proclo y Diógenes Laercio (siglo III d. C.) atribuyeron dos tipos de metodología a los platónicos. El primero es el método del análisis, en el que lo que se busca se considera como conocido, y se deducen consecuencias hasta llegar a una verdad conocida o a una contradicción; si se ha llegado a una contradicción entonces la conclusión deseada es falsa, mientras que si se ha llegado a una verdad conocida y si las etapas son reversibles se tiene una demostración. El segundo es el método de *reductio ad absurdum* o de demostración indirecta. El primer método probablemente no fue inventado por Platón, sino que quizás él subrayó su necesidad para la síntesis subsecuente, mientras que el método indirecto se le atribuye también a Hipócrates, como ya hemos indicado.

El *status* de la estructura deductiva en Platón viene expresado claramente en un pasaje de la *República* [6] en el que dice:

«Sabes bien que los estudiantes de geometría, aritmética y ciencias análogas suponen lo impar y lo par y las figuras y tres tipos de ángulos y todo lo demás en las diversas ramas de la ciencia; éstas son sus hipótesis, que se supone ellos y todo el mundo conocen, y por lo tanto no se dignan explicarlas ni a ellos mismos ni a otros, sino que comienzan por ellas y avanzan hasta que llegan al fin, y de una manera consistente, a su conclusión.»

Si admitimos que este párrafo describe fielmente la matemática de la época, entonces ya se hacían ciertamente demostraciones, pero la base axiomática quedaba implícita o podía variar algo de un matemático a otro.

Platón sostenía que era deseable una organización deductiva del conocimiento y que la tarea de la ciencia era descubrir la estructura de la naturaleza (ideal) y darle una articulación en forma de sistema

[6] Libro VI, 510.

deductivo. Platón fue el primero en sistematizar las reglas de la demostración rigurosa, y se supone que sus seguidores ordenaron los teoremas en un orden lógico. Sabemos también que en la Academia de Platón se planteó la cuestión de si un problema dado podría ser resuelto o no, sobre la base de las verdades conocidas y de las hipótesis dadas en el mismo. Hayan sido las matemáticas organizadas deductivamente a partir de axiomas explícitos por los platónicos o no, de lo que no hay duda es de que una demostración deductiva a partir de algunos principios aceptados se consideró necesaria al menos desde la época de Platón en adelante. Al insistir en esta forma de demostración los griegos rechazaban expresamente todas las reglas, procedimientos y hechos que habían sido aceptados en el «corpus» de la matemática durante miles de años antes del período griego.

¿Por qué insistieron tanto los griegos en la demostración deductiva? Dado que la inducción, observación y experimentación son fuentes vitales de conocimiento, que fueron y aún son utilizadas hoy masiva y provechosamente por las ciencias, ¿por qué prefirieron los griegos el razonamiento deductivo en matemáticas con exclusión de todos los demás métodos? Sabemos que a los griegos, o geómetras filosóficos como se llamaban, les gustaba el razonamiento y la especulación, como se pone en evidencia por sus grandes contribuciones a la filosofía, a la lógica y a la ciencia teórica, y además, los filósofos estaban interesados en alcanzar la verdad. Mientras que la inducción, la experimentación y las generalizaciones basadas en la experiencia sólo pueden dar un conocimiento probable, la deducción conduce a resultados absolutamente seguros si las premisas son correctas. En el mundo griego clásico la matemática formaba parte del cuerpo de verdades que buscaban los filósofos y así tenía que ser deductiva.

Otra razón de la preferencia griega por el método deductivo la podemos encontrar en el rechazo mostrado por la clase cultivada del período clásico griego por los asuntos prácticos. Aunque Atenas era un importante centro comercial, los negocios y otras profesiones como la medicina las realizaban los esclavos. Platón sostenía que la ocupación de los hombres libres en el comercio debería ser castigada como un crimen, y Aristóteles decía que en el estado perfecto ningún ciudadano (en oposición a los esclavos) debería practicar ningún arte mecánica. A los pensadores de tal sociedad tenían que serles extrañas la experimentación y la observación, y por lo tanto ningún resultado científico ni matemático podría derivarse de esas fuentes.

Dicho sea de paso, hay evidencia que nos muestra que en los siglos

VI y V a. C. la actitud griega hacia el trabajo, el comercio y las habilidades técnicas era muy diferente, y que la matemática encontró aplicaciones a las artes prácticas. Tales usó al parecer su matemática para mejorar la navegación, mientras que Solón, un gobernante del siglo VI, honró los oficios, y los inventores fueron muy estimados. *Sophia*, la palabra griega utilizada normalmente para expresar sabiduría y pensamiento abstracto, significaba en aquella época habilidad técnica. Fueron los pitagóricos, nos dice Proclo, los que «transformaron la matemática en una educación liberal», es decir, en una educación para hombre libres en lugar de un oficio de esclavos.

En su biografía de Marcelo, mostró Plutarco el cambio de actitud hacia artilugios tales como los instrumentos mecánicos, de la forma siguiente:

«Eudoxo y Arquitas habían sido los primeros creadores de este famoso y muy apreciado arte de la mecánica, que empleaban como elegante ilustración de las verdades geométricas y como medio de sustanciar experimentalmente, y a satisfacción de los sentidos, conclusiones demasiado complicadas para demostrarlas por medio de palabras y figuras. Como, por ejemplo, para resolver el problema, necesario a menudo en las construcciones de figuras geométricas, de dados los dos extremos, hallar los dos medios de una proporción, ambos matemáticos recurrían a la ayuda de instrumentos, adaptando a sus fines ciertas curvas y secciones de rectas. Pero con qué indignación lanzó Platón sus invectivas contra este método como la simple corrupción y anulación de lo bueno de la geometría, que volvía así vergonzosamente su espalda a los objetos incorpóreos de la inteligencia pura para recurrir a la sensación y, buscando ayuda (obtenida no sin limitaciones y pérdidas) en la materia, vino a ocurrir que la mecánica se separó de la geometría y, repudiada y despreciada por los filósofos, ocupó su lugar como arte militar.»

Esto explica el escaso desarrollo de la ciencia experimental y de la mecánica en el período clásico griego.

Haya aislado o no la investigación histórica los factores relevantes para explicar la preferencia de los griegos por el razonamiento deductivo, lo que sí sabemos con exactitud es que fueron los primeros en insistir en el razonamiento deductivo como único método de *demostración* en matemáticas. Esta exigencia ha sido característica de la matemática desde entonces, y la ha distinguido de todos los demás campos de conocimiento o investigación. Sin embargo, aún habremos de ver hasta qué punto los matemáticos posteriores permanecieron fieles a este principio.

Por lo que se refiere al contenido de la matemática, Platón y su escuela mejoraron las definiciones y se supone que demostraron nuevos teoremas de geometría plana. Además, dieron un impulso importante a la geometría del espacio. En el libro VII, sección 528 de la *República,* dice Platón que antes de estudiar la astronomía, que trata de sólidos en movimiento, se necesita una ciencia que estudie tales sólidos. Pero esta ciencia, dice, ha sido descuidada; y se queja de que los investigadores de las figuras sólidas no hayan recibido el suficiente apoyo por parte del Estado. En consecuencia, Platón y sus discípulos procedieron a estudiar la geometría del espacio, y se supone que demostraron nuevos teoremas; estudiaron las propiedades del prisma, la pirámide, el cilindro y el cono, y descubrieron que no puede haber más de cinco poliedros regulares. Los pitagóricos sabían sin duda de la existencia de tres de estos sólidos con cuatro, ocho y veinte caras triángulos equiláteros, el cubo con cuadrados y el dodecaedro con doce pentágonos, pero la demostración de que no puede haber más que cinco se debe probablemente a Teeteto.

El descubrimiento más importante quizás de la escuela platónica fue el de las secciones cónicas, atribuido por el alejandrino Eratóstenes a Menecmo, un geómetra y astrónomo que fue discípulo de Eudoxo y miembro de la Academia platónica. Aunque no sabemos con exactitud qué fue lo que llevó al descubrimiento de las secciones cónicas, se cree que resultó del estudio de los famosos problemas de construcciones. Ya hemos dicho que Hipócrates de Chíos redujo el problema de la duplicación del cubo al de hallar dos segmentos x e y tales que

$$a : x = x : y = y : 2a,$$

pero estas ecuaciones nos dicen que

$$x^2 = ay, \quad y^2 = 2ax \quad \text{y} \quad xy = 2a^2$$

y por lo tanto, usando geometría analítica podemos ver que x e y son las coordenadas del punto de intersección de dos parábolas o de una parábola y una hipérbola. Menecmo trabajó en el problema y descubrió ambas maneras de resolverlo utilizando geometría pura. Según el historiador de la matemática Otto Neugebauer, las secciones cónicas podrían haber tenido su origen en el estudio y construcción de relojes de sol.

Menecmo introdujó las secciones cónicas utilizando tres tipos de conos (fig. 3.12), con ángulos en el vértice recto, agudo y obtuso respectivamente, cortando cada uno de ellos por un plano perpendicular a una generatriz. Así pues, en esa época sólo se conocía una rama de la hipérbola.

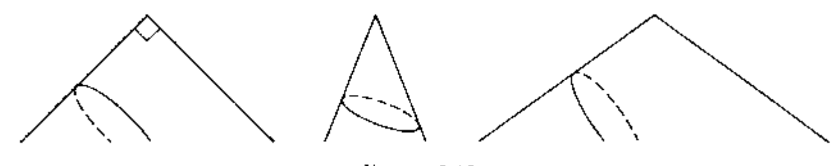


FIGURA 3.12

Entre otros estudios matemáticos debidos a los platónicos están los trabajos de Teeteto sobre los inconmensurables. Anteriormente había demostrado Teodoro de Cirene que (en nuestra notación) $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$ y otras raíces cuadradas eran irracionales. Teeteto investigó otras de tipos más complicados y las clasificó; veremos estos tipos con detalle al estudiar el libro X de los *Elementos* de Euclides. En este trabajo de Teeteto vemos cómo el sistema numérico va siendo ampliado a más irracionales, pero sólo se estudian aquellas razones inconmensurables que surgen como longitudes en construcciones geométricas. Otro discípulo de Platón, Dinostrato, mostró cómo utilizar la cuadratriz de Hipias para cuadrar el círculo y, según nos dice Pappus, Aristeo el Viejo (*c.* 320 a. C.) habría escrito una obra en cinco libros titulada *Elementos de las Secciones Cónicas.*

9. La escuela de Eudoxo

El más grande de todos los matemáticos griegos de la época clásica, superado seguramente sólo por Arquímedes en la antigüedad, fue Eudoxo, al que Eratóstenes llamó «divino». Nació en Cnido, en Asia Menor, hacia el 408 a. C., estudió con Arquitas en Tarento, viajó a Egipto, donde aprendió astronomía, y después fundó una escuela en Cyzico en el norte de Asia Menor. Hacia el 368 a. C. se unió a la escuela de Platón junto con sus discípulos, para regresar algunos años más tarde a Cnido, donde murió hacia el 355 a. C. Habiendo sido astrónomo, médico, geómetra, legislador y geógrafo, probablemente sea más conocido como creador de la primera teoría astronómica de los movimientos celestes (cap. 7).

Su primera contribución importante a la matemática fue una nueva teoría de proporciones. El descubrimiento de un número cada vez mayor de irracionales (o razones inconmensurables) hizo necesario para los griegos hacer frente a estos números; pero ¿eran realmente números? Aparecían en razonamientos geométricos mientras que los números enteros y las razones entre números enteros aparecían tanto en geometría como en el estudio general de la cantidad. Pero ¿cómo se podrían extender las demostraciones geométricas que se habían hecho para longitudes, áreas y volúmenes conmensurables a los inconmensurables?

Eudoxo introdujo la idea de magnitud continua (cap. 4, sec. 5). No se trataba de un número, sino de entidades tales como segmentos rectilíneos, ángulos, áreas, volúmenes, tiempo, etc., que podían variar, como si dijéramos, de una manera continua. Las magnitudes se oponían en esto a los números, que saltaban de un valor a otro, como del cuatro al cinco, mientras que a las magnitudes no se les asignaba ningún valor cuantitativo. Eudoxo definía entonces una razón de magnitudes y a partir de ella una proporción, es decir, una igualdad de dos razones, que cubría los casos de razones conmensurables e inconmensurables. Sin embargo, una vez más, no se utilizaba número alguno para expresar tales razones. Los conceptos de razón y proporción estaban ligados así a la geometría, como veremos al estudiar el libro V de Euclides.

Lo que consiguió así Eudoxo fue evitar los números irracionales en tanto que números, es decir, evitó darles valores numéricos a las longitudes de segmentos, tamaños de ángulos y otras magnitudes, así como a las razones de magnitudes. Mientras la teoría de Eudoxo permitió a los matemáticos griegos hacer grandes progresos en geometría, suministrándoles los fundamentos lógicos necesarios para las razones inconmensurables, también tuvo varias consecuencias desafortunadas.

Por mencionar una, forzó una nítida separación entre número y geometría, dado que únicamente la geometría podía manejar las razones inconmensurables, pero también hizo de los matemáticos geómetras, y la geometría iba a convertirse en la base de casi toda la matemática rigurosa durante los dos mil años siguientes. Nosotros decimos aún x^2, «x cuadrado» y x^3, «x cubo» en lugar de, digamos, x segunda o x tercera, debido a que las magnitudes x^2, y x^3 sólo tenían un significado geométrico para los griegos.

La solución de Eudoxo al problema de cómo tratar las magnitudes

inconmensurables o los números irracionales invirtió, de hecho, el punto de vista de la matemática griega anterior. Los pitagóricos primitivos habían puesto ciertamente el énfasis en el número como concepto fundamental, y Arquitas de Tarento, maestro de Eudoxo, afirmaba que sólo la aritmética y no la geometría podía dar demostraciones satisfactorias. Sin embargo, al volver a la geometría para manejar los números irracionales, los griegos abandonaron el álgebra y los números irracionales como tales. Pero ¿qué es lo que hicieron para resolver ecuaciones cuadráticas, donde las soluciones son frecuentemente números irracionales?, y ¿de qué manera trataron el sencillo problema de hallar el área de un rectángulo de lados inconmensurables? La respuesta a ambas preguntas es la de que transformaron la mayor parte del álgebra en geometría, en un proceso que analizaremos en el capítulo siguiente. La representación geométrica de los irracionales y de las operaciones con ellos no era práctica, evidentemente. Puede resultar lógicamente satisfactorio pensar en $\sqrt{2} \cdot \sqrt{3}$ como el área de un rectángulo, pero si se necesita saber el producto para comprar moqueta, por ejemplo, no dará resultado.

Aunque los griegos dedicaron sus mayores esfuerzos en matemáticas a la geometría, no hay que olvidar que los números enteros y las razones entre ellos siguieron siendo conceptos perfectamente aceptables. Este campo de la matemática, como veremos, aparece organizado deductivamente en los libros VII, VIII y IX de los *Elementos* de Euclides; el material en cuestión cubre esencialmente lo que llamamos teoría de números o estudio de las propiedades de los enteros.

La siguiente pregunta se plantea de manera natural: ¿Qué hicieron los griegos con la necesidad de los números en la investigación científica, así como en el comercio y otros asuntos prácticos? Por un lado, la ciencia griega clásica fue cualitativa, como veremos. En cuanto a los usos prácticos de los números, ya hemos dicho que los intelectuales de la época se limitaron a las actividades filosóficas y científicas y no se ocuparon del comercio ni de los oficios; el hombre cultivado no se interesaba por los problemas prácticos. Pero uno puede pensar en todos los rectángulos de la geometría sin referirse para nada a las dimensiones concretas de ninguno de ellos. El pensamiento matemático se vio así separado de las necesidades prácticas, y los matemáticos no encontraron motivación para mejorar las técnicas aritméticas y algebraicas. Cuando las barreras ente las clases cultivadas y los esclavos se hicieron menos estrictas en el período alejandrino (del 300 a. C. al 600 d. C. aproximadamente) y los hombres cultos se interesa-

ron por los asuntos prácticos, el énfasis se desplazó al conocimiento cuantitativo y al desarrollo de la aritmética y el álgebra.

Volviendo a las contribuciones de Eudoxo, también se debe a él el poderoso método griego para hallar áreas y volúmenes de figuras curvilíneas que nosotros llamamos método de exhausción. Estudiaremos este método y sus aplicaciones, tal como las presenta Euclides, más adelante. Se trata realmente de la primera etapa en la historia del cálculo infinitesimal, pero no utiliza una teoría de límites explícita. Con su ayuda demostró Eudoxo, por ejemplo, que las áreas de dos círculos son entre sí como los cuadrados de sus radios, que los volúmenes de dos esferas son entre sí como los cubos de sus radios, que el volumen de una pirámide es un tercio del volumen del prisma con la misma base y altura, y que el volumen de un cono es un tercio del cilindro correspondiente.

Siempre puede encontrarse un motivo u otro para atribuir a cualquier escuela, desde Tales en adelante, el haber introducido la organización deductiva de la matemática, pero es incuestionable, sin embargo, que la obra de Eudoxo estableció la organización deductiva *sobre la base de unos axiomas explícitos.* La razón para ello fue sin duda la necesidad de entender y operar con razones inconmensurables. Dado que Eudoxo abordó la tarea de construir la base lógica precisa para estas razones, es lo más verosímil que viera la necesidad de formular axiomas y deducir consecuencias una por una de manera que no se cometieran errores con estas magnitudes extrañas y conflictivas. Esta necesidad de manejar razones inconmensurables vino también a reforzar, sin duda, la decisión anterior de apoyarse exclusivamente en el razonamiento deductivo en las demostraciones.

Como los griegos buscaban verdades y habían decidido utilizar las demostraciones deductivas, tenían que basarse en axiomas que fueran ellos mismos verdaderos, y encontraron en efecto afirmaciones cuya veracidad era evidente para ellos, aunque las justificaciones dadas para aceptar los axiomas como verdades indiscutibles fueran diversas. Casi todos los griegos creían que la mente era capaz de reconocer estas verdades y Platón, en particular, aplicó su teoría de la *anamnesis,* según la cual hemos tenido ya una experiencia directa de las verdades en un período de existencia como almas en otro mundo antes de venir a la tierra, y no tenemos más que recordar esta experiencia para saber que estas verdades incluyen a los axiomas de la geometría; no es necesaria ninguna experiencia en la tierra. Algunos historiadores pretenden ver en las teorías de Platón y Proclo la idea de que puede

haber alguna arbitrariedad en los axiomas, con tal solamente de que sean claros y verdaderos en la mente del matemático individual. Lo importante es razonar deductivamente sobre la base de los axiomas elegidos. Aristóteles tenía mucho que decir sobre los axiomas y pasamos ahora a exponer sus puntos de vista.

10. Aristóteles y su escuela

Aristóteles (384-322 a. C.) nació en Estagira, ciudad de Macedonia. Durante 20 años fue discípulo de Platón y durante otros 3 años, del 343 al 340 a. C., fue tutor de Alejandro Magno. El año 335 a. C. fundó su propia escuela, el Liceo, con un jardín, un aula y un altar a las Musas.

Aristóteles escribió sobre mecánica, física, matemáticas, lógica, meteorología, botánica, psicología, zoología, ética, literatura, metafísica, economía y muchos otros temas. No hay ningún libro dedicado exclusivamente a la matemática, pero en diversos lugares aparecen discusiones sobre la materia, que utiliza como ejemplos en muchos contextos.

Aristóteles consideraba a las ciencias clasificadas en tres tipos: teóricas, productivas y prácticas. Las teóricas, que son las que buscan la verdad, son la matemática, la física (óptica, armonía y astronomía), y la metafísica; de ellas la más exacta es la matemática. Las ciencias productivas son en realidad las artes, y las prácticas, como por ejemplo la ética y la política, tratan de regular las acciones humanas. En las ciencias teóricas la lógica es previa a los diversos temas incluidos en ellas, y el metafísico discute y explica lo que el matemático y el filósofo natural (o científico) toma como dado, por ejemplo el ser o realidad de la materia y el tipo de los axiomas.

Aunque Aristóteles no contribuyó con resultados matemáticos nuevos de importancia (algunos teoremas de Euclides se le atribuyen, sin embargo), sus teorías sobre la naturaleza de la matemática y sus relaciones con el mundo físico ejercieron una gran influencia. Mientras Platón creía que había un mundo independiente y eterno de las ideas, que constituía la realidad del universo y del que formaban parte los conceptos matemáticos, Aristóteles atribuía este papel a la materia o substancia concreta. Sin embargo, también llegó a poner énfasis en las ideas, es decir en las esencias universales de los objetos físicos, tales como dureza, blandura, gravedad, ligereza, esfericidad, frialdad y

calor. Los números y las formas geométricas eran también propiedades de los objetos reales; se reconocían por abstracción pero pertenecían en realidad a los objetos mismos. Así, la matemática trabaja con conceptos abstractos que se derivan de propiedades de los cuerpos físicos.

Aristóteles discute también el concepto de definición. Su idea de definición es la moderna y la denomina un nombre para una colección de palabras señalando también que una definición correcta debe estar expresada en términos de algo previo a la cosa definida. Así, por ejemplo, critica la definición «un punto es aquello que no tiene partes», porque las palabras «aquello que» no dicen a qué se refieren, excepto posiblemente a «punto», y por lo tanto la definición no sería correcta. Reconoce, evidentemente, la necesidad de términos indefinidos, puesto que debe haber un punto de partida para la serie de definiciones, pero los matemáticos posteriores olvidaron esta necesidad hasta finales del siglo XIX.

Advierte también (como había hecho anteriormente Platón, según Plutarco) que una definición nos dice lo que es una cosa, pero no que la cosa misma exista. La existencia de las cosas definidas tiene que demostrarse excepto en el caso de unas pocas cosas primarias tales como el punto y la recta, cuya existencia se supone en los primeros principios o axiomas. Podemos definir un cuadrado pero tal figura puede no existir, es decir, las propiedades exigidas en la definición pueden ser incompatibles. Leibniz puso el ejemplo de un poliedro regular de 10 caras; uno puede definir naturalmente tal figura, pero no existe. Si no se comprueba que esta figura existe, y procedemos a demostrar teoremas acerca de ella, los resultados no tendrán sentido. El método de demostrar la existencia que adoptaron Aristóteles y Euclides fue el de la construcción. Los tres primeros axiomas de los *Elementos* de Euclides garantizan la construcción de rectas y circunferencias; todos los conceptos matemáticos restantes han de ser construidos para establecer su existencia. Así, los trisectores de ángulos, aunque sean evidentemente definibles, no son constructibles con rectas y circunferencias y por lo tanto no podían admitirse en la geometría griega.

Aristóteles se ocupa también de los principios básicos de la matemática, distinguiendo entre los axiomas o nociones comunes, que son verdades comunes a todas las ciencias, y los postulados, que son primeros principios aceptables para una ciencia concreta. Entre los axiomas incluye los principios lógicos, tales como la ley de contradic-

ción, la ley del tercio excluso, el axioma de que si se suman o restan cosas iguales de otras iguales los resultados son iguales, y otros principios análogos. Los postulados no necesitan ser auto-evidentes sino que su verdad debe venir garantizada por las consecuencias que se derivan de ellos. La colección de axiomas y postulados ha de ser lo más reducida posible, con tal de que permitan demostrar todos los resultados. Aunque, como veremos, Euclides utiliza la distinción de Aristóteles entre nociones comunes y postulados, todos los matemáticos hasta principios del XIX pasaron por alto esta distinción y trataron los axiomas y los postulados como igualmente auto-evidentes. Según Aristóteles, los axiomas se obtienen de la observación de los objetos físicos, de los que son generalizaciones aprehendidas de modo inmediato. Tanto él como sus seguidores dieron muchas definiciones y axiomas o mejoraron otros anteriores y algunas de las versiones aristotélicas las incluye directamente Euclides.

Aristóteles discute los problemas fundamentales acerca de las relaciones entre puntos y rectas. Un punto, dice, es indivisible y tiene posición; pero entonces ninguna acumulación de puntos, por muchos que incluyera, podría darnos algo divisible, mientras que una recta es desde luego una magnitud divisible. Por lo tanto los puntos no pueden construir nada continuo como una recta, pues un punto no puede ser continuo con otro punto. Un punto, añade, es como el ahora en el tiempo; el ahora es indivisible y no una parte del tiempo. Un punto puede ser un comienzo, un final o un divisor en un segmento pero no es parte de él ni de ninguna magnitud. Solamente por *movimiento* puede un punto generar una recta y ser así origen de la magnitud. También afirma que, puesto que un punto no tiene longitud, si una recta estuviera compuesta de puntos tampoco tendría longitud y, análogamente, si el tiempo estuviera compuesto de instantes no habría ningún intervalo de tiempo. Su definición de continuidad, propiedad que posee una recta, es la siguiente: una cosa es continua cuando los límites en los que se tocan dos partes sucesivas cualesquiera son uno y el mismo y están, como la palabra misma continuo implica, juntos. En realidad hace diversas afirmaciones sobre las magnitudes continuas que no concuerdan unas con otras. El núcleo de su teoría, sin embargo, es que los puntos y los números son cantidades discretas y hay que distinguirlas de las magnitudes continuas de la geometría; no hay continuo en la aritmética. En cuanto a la relación entre estos dos campos, considera a la aritmética (es decir, a la teoría de números) como más exacta, porque los números se prestan

más fácilmente a la abstracción que los conceptos geométricos. También considera a la aritmética como previa a la geometría, porque el número 3 es necesario para considerar un triángulo.

Al discutir el infinito hace Aristóteles una distinción, importante aún hoy, entre el infinito potencial y el infinito actual. La edad de la tierra, si es que tuvo un comienzo, es potencialmente infinita pero en ningún instante es actualmente infinita. Según él, sólo existe el infinito potencial. Los enteros positivos, concede, son potencialmente infinitos porque siempre podemos añadir 1 a cualquier número y obtener otro distinto, pero el conjunto infinito, como tal, no existe. La mayor parte de las magnitudes, incluso, no pueden ser ni siquiera potencialmente infinitas, porque si se añadieran de una manera indefinida podrían exceder los límites del universo. El espacio, sin embargo, sí es potencialmente infinito en el sentido de que puede ser subdividido indefinidamente, y el tiempo es potencialmente infinito en los dos sentidos.

Uno de los logros más importantes de Aristóteles fue la fundamentación de la ciencia de la lógica. Los griegos habían hecho ya el trabajo básico para fundar la lógica al producir razonamientos matemáticos correctos, pero correspondió a Aristóteles codificar y sistematizar las leyes que siguen estos razonamientos en una disciplina separada. Los escritos de Aristóteles dejan muy claro que derivó la lógica de la matemática. Los principios básicos de su lógica —la ley de contradicción, que afirma que una proposición no puede ser a la vez verdadera y falsa, y la ley de tercio excluso, que afirma que una proposición debe ser verdadera o falsa— están en el centro mismo del método de demostración indirecto en matemáticas; por otra parte, Aristóteles utiliza abundantes ejemplos matemáticos tomados de textos contemporáneos para ilustrar sus principios de razonamiento. Esta lógica aristotélica permaneció insuperada hasta el siglo XIX.

A pesar de que la lógica se derivó de la matemática, finalmente fue considerada como independiente de y previa a la matemática, y aplicable a todos los razonamientos. Incluso Aristóteles, como ya hemos hecho notar, consideró a la lógica como preliminar a las ciencias y a la filosofía, y concretamente en matemáticas subrayó la demostración deductiva como la única base para establecer los hechos. Para Platón, que creía que las verdades matemáticas preexistían o existían en un mundo independiente del hombre, el razonamiento no era la garantía de la corrección de los teoremas; el poder de la

lógica jugaba sólo un papel secundario, haciendo explícito, por así decirlo, lo que ya se sabía que era verdadero.

Un miembro de la escuela aristotélica especialmente digno de mención es Eudemo de Rodas, que vivió a finales del siglo IV a. C. y fue el autor del «Sumario de Eudemo» citado por Proclo y por Simplicio. Como ya dijimos, Eudemo escribió historias de la aritmética, de la geometría y de la astronomía. Se trata pues del primer historiador de la ciencia que conocemos, pero lo que es más importante es que los conocimientos ya existentes en su época fueran lo suficientemente amplios como para merecer ser historiados.

El último de los autores del período clásico que vamos a mencionar aquí es Autólico de Pitania, astrónomo y geómetra que floreció hacia el 310 a. C. No fue miembro de la escuela de Platón ni de la de Aristóteles, aunque fue maestro de uno de los sucesores de Platón. De tres libros que escribió nos han llegado dos; son los libros griegos más antiguos que conocemos completos, aunque sólo a través de manuscritos que presumiblemente son copias de las obras de Autólico. Estos libros, *Sobre la Esfera en Movimiento* y *Sobre Salidas y Puestas* fueron incluidos más tarde en una colección llamada *Pequeña Astronomía* (para distinguirla de la posterior *Gran Colección* o *Almagesto* de Ptolomeo). *Sobre la Esfera en Movimiento* trata de los círculos meridianos, de los círculos máximos en general, y de lo que llamaríamos paralelos de latitud, así como de las áreas visible e invisible producidas por una fuente luminosa distante sobre una esfera en rotación, tal como el sol sobre la tierra. El libro presupone teoremas de geometría esférica que debían conocer, por lo tanto, los griegos de la época. El segundo libro de Autólico sobre la salida y puesta de estrellas corresponde a la astronomía de observación.

La forma del libro sobre la esfera en movimiento es importante; los puntos de las figuras vienen representados por letras y las proposiciones están ordenadas lógicamente. Primero se formula la proposición en general, después se repite, pero con referencia explícita a la figura y finalmente se da la demostración. Este es ya el estilo que usa Euclides.

Bibliografía

Apostle, H. G.: *Aristotle's Philosophy of Mathematics*, University of Chicago Press, 1952.

Ball, W. W. Rouse: *A Short Account of the History of Mathematics*, Dover (reprint), 1960, caps. 2-3.

Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Brumbaugh, Robert S.: *Plato's Mathematical Imagination*, Indiana University Press, 1954.

Gomperz, Theodor: *Greek Thinkers*, 4 vols., John Murray, 1920.

Guthrie, W. K. C.: *Historia de la filosofía griega*, Madrid, Gredos, 1984.

Hamilton, Edith: *The Greek Way to Western Civilization*, New American Library, 1948.

Heath, Thomas L.: *A History of Greek Mathematics*, Oxford University Press, 1921, vol. 1.

—: *A Manual of Greek Mathematics*, Dover (reprint), 1963, caps. 4-9.

—: *Mathematics in Aristotle*, Oxford University Press, 1949.

Jaeger, Werner: *Paideia: los ideales de la cultura griega*, México, FCE, 1990.

Lasserre, François: *The Birth of Mathematics in the Age of Plato*, American Research Council, 1964.

Maziarz, Edward A., y Thomas Greenwood: *Greek Mathematical Philosophy*, F. Unger, 1968.

Sarton, George: *A History of Science*, Harvard University Press, 1952, vol. 1, caps. 7, 11, 16, 17 y 20.

Scott, J. F.: *A History of Mathematics*, Taylor and Francis, 1958, cap. 2.

Smith, David Eugene: *History of Mathematics*, Dover (reprint), 1958, vol. 1, cap. 3.

Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, caps. 4-6.

Wedberg, Anders: *Plato's Philosophy of Mathematics*, Almqvist and Wiksell, 1955.

Capítulo 4
EUCLIDES Y APOLONIO

> Aprendimos de los pioneros en esta ciencia a no atender a meras imágenes plausibles cuando se trata de los razonamientos que deben presentarse en nuestra doctrina geométrica.
>
> PROCLO

1. Introducción

Lo más importante de la obra matemática que realizaron los autores del período clásico ha llegado afortunadamente hasta nosotros, en los escritos de Euclides y Apolonio. Cronológicamente, ambos pertenecen al segundo gran período de la historia griega, el helenístico o alejandrino. Sabemos con certeza, gracias a un párrafo del *Comentario* de Proclo, que Euclides vivió y enseñó en Alejandría en torno al año 300 a. C., aunque probablemente se educara en la Academia de Platón; y esto es todo cuanto conocemos de su vida. Apolonio murió en el año 190 a. C., de modo que toda su vida cae claramente dentro del período helenístico. Es habitual, sin embargo, situar su obra en el período clásico, ya que sus libros dan cuenta de lo producido en tal época. De hecho, Euclides estructuró los descubrimientos dispares de los griegos clásicos, como puede comprobarse comparando el contenido de sus libros con los fragmentos que nos han llegado de trabajos más antiguos; constituyen así los *Elementos* tanto una historia matemática de la época precedente como el desarrollo lógico de una teoría. La obra de Apolonio se sitúa generalmente en el período alejandrino que le corresponde, pero el espíritu y el

contenido de su principal trabajo, las *Secciones Cónicas*, son del período clásico. El mismo Apolonio reconoció que los cuatro primeros libros de los ocho que lo forman constituyen una revisión de los trabajos perdidos de Euclides sobre el mismo tema. Pappus menciona que Apolonio pasó largo tiempo con los discípulos de Euclides en Alejandría, lo que explica su familiaridad con la obra de este último. La discusión que haremos más adelante sobre las características del período alejandrino justificará, a nuestro parecer, la inclusión de Apolonio en el período clásico.

2. El marco de los *Elementos* de Euclides

Los *Elementos* son sin duda la obra más famosa de Euclides. Pese al escaso conocimiento que poseemos del período clásico, cabe señalar las principales fuentes del material contenido en ellos: aparte de los discípulos de Platón con quienes estudió Euclides, y a quienes debe mucho, sin duda, Proclo afirma que introdujo en sus *Elementos* muchos de los teoremas de Eudoxo, perfeccionó teoremas de Teeteto y proporcionó demostraciones irrefutables de muchos resultados insuficientemente demostrados por sus predecesores.

A Euclides se debe la elección del sistema de axiomas, la ordenación de los teoremas y la tersura y rigor de las demostraciones, muchas de ellas suyas, sin duda. Su forma de presentar éstas, sin embargo, había sido ya empleada por Autólico (ver cap. 3, sec. 9) y seguramente por otros de sus predecesores. Independientemente de la cuestión de cuánto haya de original en sus *Elementos* y cuánto pudo haber recogido de textos anteriores u otras fuentes, Euclides fue sin duda un gran matemático, como lo prueban sus otros escritos. Proclo señala que los *Elementos* eran muy apreciados en Grecia, e indica como prueba el gran número de comentarios a que habían dado lugar; entre los más importantes cabe citar los de Herón (*c.* 100 a. C.-*c.* 100 d. C.), Porfirio (siglo III) y Pappus (finales del mismo siglo). Presumiblemente su calidad les permitió reemplazar a los libros que sobre el mismo asunto se cree que escribieron Hipócrates de Chíos y los platónicos León y Teudio.

No contamos con manuscritos del propio Euclides, y sus escritos han tenido que ser reconstruidos a partir de las numerosas recensiones, comentarios y notas de otros autores. Todas las ediciones en lengua inglesa y en latín de los *Elementos* se han realizado a partir de

manuscritos griegos; la recensión de Teón de Alejandría (fines del siglo IV), copias de ésta, versiones escritas de las lecciones de Teón, y un manuscrito griego del siglo X que François Peyrard (1760-1822) halló en la Biblioteca Vaticana, y que es una copia de una edición de Euclides anterior a la de Teón. Los historiadores J. L. Heiberg y Thomas L. Heath han utilizado principalmente este manuscrito en su estudio sobre Euclides, comparándolo, claro está, con los restantes manuscritos y comentarios disponibles. También existen versiones y comentarios árabes, basados al parecer en manuscritos griegos que se han perdido, y que han servido para precisar qué había y qué no en los *Elementos* escritos por Euclides; pero estas versiones árabes son en cualquier caso inferiores a los manuscritos griegos. Al apoyarse en tantas fuentes, la reconstrucción de los *Elementos* deja margen para la duda sobre algunas cuestiones. En particular, no sabemos con qué propósito fueron escritos; hay quienes los consideran un tratado para matemáticos formados, y quienes piensan que se trata de un texto para estudiantes. Proclo parece inclinarse por esta última opción.

Dadas la longitud y la incomparable importancia histórica de esta obra, dedicaremos varias secciones del presente capítulo a un repaso y comentario de su contenido, que quizá nos sorprenda un poco al compararlo con la «geometría euclídea» aprendida en la enseñanza media o superior. Las versiones más ampliamente difundidas en nuestro tiempo se basan en la modificación que Legendre realizó de los trabajos de Euclides, empleando una pizca de álgebra ajena a los *Elementos*, sin que por ello, como veremos, se altere el material geométrico.

3. Las definiciones y axiomas de los *Elementos*

Los *Elementos* constan de trece libros. En algunas ediciones se han incluido otros dos, debidos probablemente a otros autores. El libro I comienza con las definiciones de los conceptos que se utilizarán en la primera parte de la obra. Copiaremos aquí sólo las más importantes, numerándolas de acuerdo con la edición de Heath [1]:

---

[1] T. L. Heath: *The Thirteen Books of Euclid's Elements*, Dover (reimpresión), 1956, en 3 volúmenes.

Definiciones

1. Un punto es lo que no tiene partes.

2. Una línea es una longitud sin anchura.
   La palabra línea significa curva.

3. Los extremos de una línea son puntos.
   Esta definición establece que una línea o curva siempre tiene longitud finita; en los *Elementos* no aparecen curvas que se extiendan hasta el infinito.

4. Una línea recta es aquella que yace por igual sobre sus puntos.
   De acuerdo con la definición 3, la línea recta de Euclides es nuestro segmento. Se cree que esta definición pudo ser sugerida por el nivel que se usa en albañilería.

5. Una superficie es lo que sólo tiene longitud y anchura.

6. Los extremos de una superficie son líneas.
   Una superficie, por tanto, es también una figura acotada.

7. Una superficie plana es la que yace por igual sobre sus líneas rectas.

15. Un círculo es una figura plana rodeada por una línea tal que todas las rectas que inciden sobre ella desde cierto punto interior a la figura son iguales entre sí.

16. Ese punto se llama centro del círculo.

17. Un diámetro del círculo es cualquier recta que pasa por el centro y cuyos extremos están en la circunferencia (no definida explícitamente) del círculo. Tal recta divide en dos partes iguales al círculo.

23. Rectas paralelas son aquellas que, estando en el mismo plano, no se encuentran cuando se prolongan indefinidamente en ambas direcciones.

Estas definiciones liminares vienen cargadas de conceptos no definidos y no convienen por tanto a ningún propósito lógico. Puede que Euclides no se apercibiera de que los conceptos iniciales deben quedar sin definición, lo que le habría llevado a explicar ingenuamente su significado en términos de conceptos físicos. Algunos comentaristas afirman que, aun siendo consciente de que las definiciones no eran lógicamente útiles, quiso explicar lo que sus términos representaban

intuitivamente, de manera que sus lectores quedaran convencidos de que los axiomas y postulados eran aplicables a esos conceptos.

A continuación presenta cinco postulados y cinco nociones comunes (a las que Proclo llama axiomas). Asume la distinción ya indicada por Aristóteles de que las nociones comunes son verdades aplicables a cualquier ciencia, mientras que los postulados se aplican solamente a la geometría. Como ya señalamos en su momento, Aristóteles decía que no se precisa la certeza de que los postulados sean verdaderos, y que su veracidad se contrastaría al confrontar con la realidad los resultados de ellos deducidos. Proclo incluso habla del carácter hipotético de toda matemática, que sólo deduce lo que se sigue de las suposiciones iniciales, sean éstas verdaderas o no. Cabe pensar que Euclides compartiera el punto de vista de Aristóteles con respecto a la veracidad de los postulados. No obstante, en el desarrollo ulterior de las matemáticas, al menos hasta el advenimiento de las geometrías no euclídeas, tanto los postulados como las nociones comunes fueron aceptados como verdades incuestionables.

Euclides postula lo siguiente:

Postulados

1. (Es posible) trazar una línea recta desde cualquier punto a cualquier otro.

2. (Es posible) prolongar continuamente en línea recta una recta dada.

3. (Es posible) trazar un círculo con cualquier centro y distancia (radio).

4. Que todos los ángulos rectos son iguales entre sí.

5. Que si una recta incide sobre otras dos formando del mismo lado ángulos internos menores que dos rectos, al prolongarlas indefinidamente se encontrarán por el lado en que los ángulos sean menores que dos rectos.

Nociones comunes

1. Cosas que sean iguales a una misma cosa son también iguales entre sí.

2. Si a cosas iguales se suman cosas iguales, los totales son iguales.

3. Si a cosas iguales se restan cosas iguales, los restos son iguales.

4. Cosas que encajen cada una en la otra son iguales entre sí.

5. El todo es mayor que la parte.

Euclides no supone ingenuamente que los conceptos definidos existan o sean consistentes; como había señalado Aristóteles, se puede definir algo cuyas propiedades sean incompatibles. Los tres primeros postulados, que declaran la posibilidad de construir rectas y círculos, son asertos de existencia para esas entidades. A lo largo del libro I, Euclides prueba, construyéndolas, la existencia de las restantes, exceptuado el plano.

Presupone que la recta del postulado 1 es única; esta suposición está implícita en la proposición 4 del libro I, aunque habría sido mejor explicitarla. Del mismo modo, supone que la prolongación del postulado 2 es única, explícitamente en la proposición 1 del libro XI, e inconscientemente desde el mismo comienzo del libro I.

El postulado 5 se debe al propio Euclides; es una muestra de su genio haber reconocido su necesidad. Muchos griegos objetaron este postulado, considerándolo falto de evidencia, en comparación con los anteriores. Los intentos de probarlo a partir de los restantes axiomas y postulados, que comenzaron según Proclo en vida misma de Euclides, fracasaron. La historia detallada de tales esfuerzos se verá más adelante, en relación con la discusión sobre geometrías no euclídeas.

En cuanto a las nociones comunes, hay diferentes opiniones sobre cuáles aparecían realmente en el escrito original de Euclides. La cuarta, que constituye la base de las pruebas mediante superposición (congruencia) es de carácter geométrico, y debería ser un postulado. Euclides la utiliza en las proposiciones 4 y 8 del libro I, aunque diríase que de mala gana; podría haber hecho uso de ella en la demostración de la proposición 26 (*a.l.a.* = *a.l.a.* y *l.a.a.* = *l.a.a.*) y en cambio presenta una prueba más larga. Probablemente conocía el método por los trabajos de anteriores geómetras, y no supo cómo evitarlo. Pappus y otros, que encontraron inadecuado el conjunto de axiomas propuesto por Euclides, añadieron alguno más.

4. Los libros I a IV de los *Elementos*

Los libros I a IV tratan sobre las propiedades básicas de figuras rectilíneas y círculos. El libro I contiene los acostumbrados teoremas

sobre congruencia, paralelismo, el teorema de Pitágoras, figuras equivalentes (de igual área) y paralelogramos. Todas las figuras son rectilíneas, esto es, formadas por segmentos de recta. De especial interés son los siguientes teoremas (la versión no es literal):

Proposición 1. Construcción de un triángulo equilátero sobre un segmento dado.

La demostración es simple. Se construye un círculo tomando *A* como centro y *AB* como radio (fig. 4.1), y otro con *B* como centro y *BA* como radio. Sea *C* el punto de intersección. Entonces *ABC* es el triángulo buscado.

Proposición 2. Situar en un punto dado (como extremo) una línea recta igual a otra dada.

Podría pensarse que el postulado 3 permite hacerlo inmediatamente. Pero eso significaría que el compás mantiene su abertura cuando se lleva al punto que se quiere tomar como extremo. Euclides, en cambio, supone un compás que sólo mantiene su rigidez al trazar un círculo determinado, sin levantarlo del papel, y presenta una demostración más complicada.

Proposición 4. Si dos triángulos tienen cada uno de ellos dos lados y el ángulo que comprenden iguales a los del otro, entonces son congruentes.

La prueba se hace llevando un triángulo sobre el otro, y mostrando que deben coincidir.

Proposición 5. Los ángulos de la base de un triángulo isósceles son iguales.

La demostración es mejor que la que puede encontrarse en muchos textos elementales, que emplea la bisectriz del ángulo *A* (fig. 4.2), cuya existencia se deduce precisamente de esta proposición. Euclides extiende *AB* hasta *F* y *AC* hasta *G*, de manera que $BF = CG$. Entonces $\triangle AFC \cong \triangle AGB$, y por tanto $FC = GB$, $\measuredangle ACF = \measuredangle ABG$ y $\measuredangle 3 = \measuredangle 4$. De esto se deduce que $\triangle CBF \cong \triangle BCG$ y por tanto $\measuredangle 5 = \measuredangle 6$, y $\measuredangle 1 = \measuredangle 2$. Pappus prueba el teorema considerando el triángulo dado como $\triangle ABC$ y como $\triangle ACB$, lo que le permite utilizar la proposición 4 y deducir que los ángulos de la base son iguales.

Proposición 16. Un ángulo exterior de un triángulo es mayor que cualquiera de los dos ángulos internos opuestos.

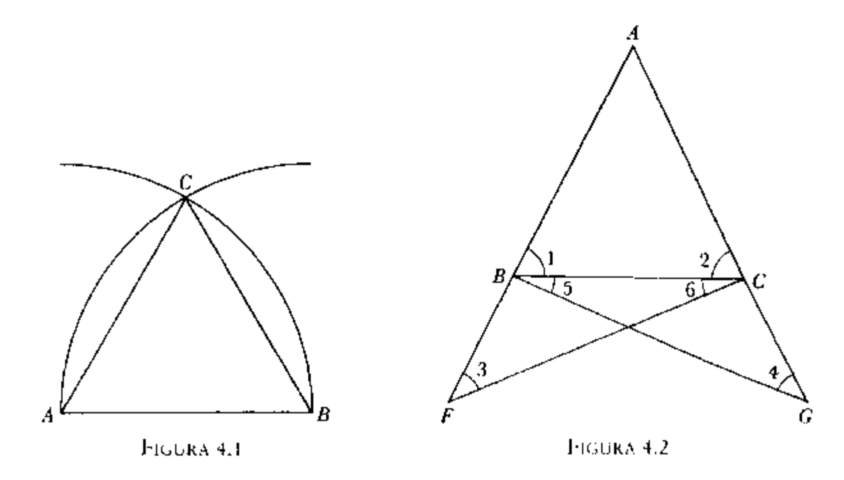


Figura 4.1

Figura 4.2

La prueba (fig. 4.3) requiere una recta indefinidamente prolongable, ya que en ella se extiende AE una longitud igual hasta F, y ha de ser posible hacer esto.

Proposición 20. La suma de dos lados cualesquiera de un triángulo es mayor que el tercer lado.

Este teorema es lo que más se parece en geometría euclídea al hecho de que la línea recta es la distancia más corta entre dos puntos.

Proposición 27. Si una recta incide sobre otras dos formando ángulos alternos iguales, esas dos rectas serán paralelas entre sí.

La prueba aportada consiste en suponer que las rectas se cortan, de lo que se deriva una contradicción con la proposición sobre el ángulo externo de un triángulo. El teorema establece la existencia de al menos una recta paralela a otra dada, pasando por un punto también dado.

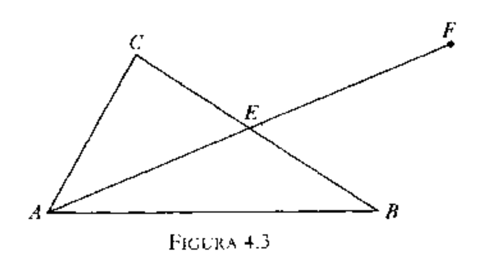


Figura 4.3

Proposición 29. Una recta que incide sobre dos paralelas forma ángulos alternos iguales entre sí, siendo cada ángulo externo igual al interno opuesto (los ángulos correspondientes son iguales), y la suma de los ángulos internos del mismo lado es igual a dos rectos.

La demostración (fig. 4.4) supone que ∢1 ≠ ∢2. Si el mayor es ∢2, sumando ∢4 a ambos, ∢2 + ∢4 > ∢1 + ∢4, lo que implica que ∢1 + ∢4 es menor que dos rectos. Pero el postulado de las paralelas, que es utilizando aquí por primera vez, implicaría que las rectas *AB* y *CD*, que por hipótesis son paralelas, se encuentran en algún punto.

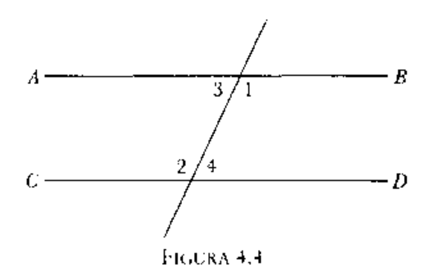


FIGURA 4.4

Proposición 44. Construir sobre una recta dada, y con un ángulo rectilíneo dado, un paralelogramo equivalente a un triángulo dado.

A partir de un triángulo *C*, un ángulo *D* y un segmento *AB* (fig. 4.5), la proposición afirma la posibilidad de construir un paralelogramo que tenga *AB* como lado, *D* como ángulo, y cuya área sea igual a la de *C*. No expondremos la demostración de Euclides, que depende de otras proposiciones anteriores. Lo que importa señalar aquí es que se trata del primero de los problemas de aplicación de áreas, teoría que Eudemo (según cuenta Proclo) atribuía a los pitagóricos. En este caso se aplica (exactamente) un área al segmento *AB*. En

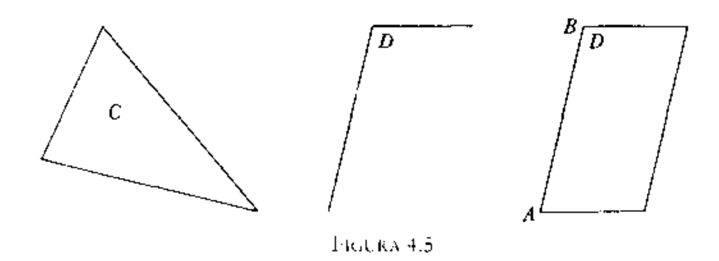


FIGURA 4.5

segundo lugar, se trata de un ejemplo de transformación de un área en otra. Por último, en el caso especial de que D sea un ángulo recto, el paralelogramo debe ser un rectángulo. Entonces, el área del triángulo dado y AB pueden considerarse como cantidades dadas, y el otro lado del rectángulo será el cociente entre el área de C y AB, habiéndose llevado a cabo la división geométricamente; este teorema es un ejemplo de álgebra geométrica.

Proposición 47. En los triángulos rectángulos el cuadrado del lado opuesto al ángulo recto es igual a la suma de los cuadrados de los lados que lo forman.

Tenemos aquí el teorema de Pitágoras. La prueba se lleva a cabo por medio de áreas, como en muchos textos escolares. Se muestra (fig. 4.6) que $\triangle ABD \cong \triangle FBC$, que el rectángulo $BL = 2\triangle ABD$, y el rectángulo $GB = 2\triangle FBC$. En consecuencia, el rectángulo BL es igual al cuadrado GB, y el rectángulo CL es igual al cuadrado AK.

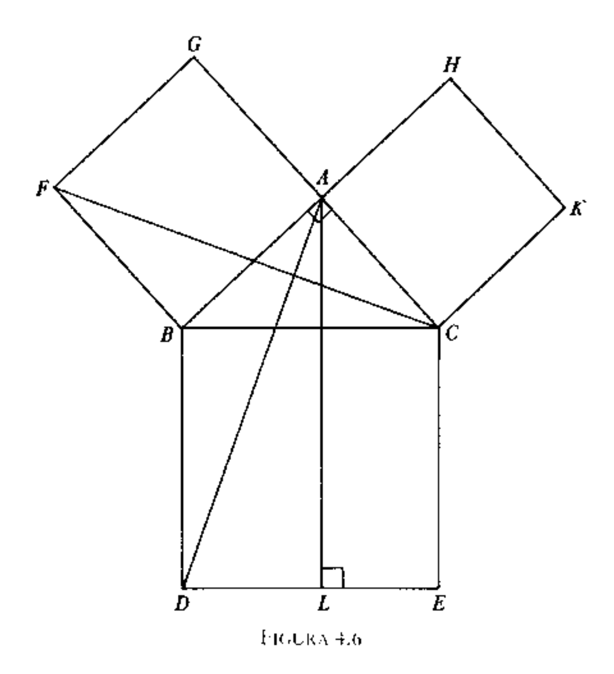


FIGURA 4.6

El teorema también muestra cómo obtener un cuadrado cuya área sea igual a la suma de dos cuadrados dados, es decir, cómo hallar un x tal que $x^2 = a^2 + b^2$, siendo así otro ejemplo de álgebra geométrica.

Proposición 48. Si en un triángulo el cuadrado de un lado es igual a la suma de los cuadrados de los otros dos lados, el ángulo que éstos forman es recto.

Esta proposición es la recíproca del teorema de Pitágoras. La demostración de Euclides (fig. 4.7) consiste en trazar un segmento AD perpendicular a AC e igual a AB. Por hipótesis,

$$AB^2 + AC^2 = BC^2,$$

y por ser rectángulo el triángulo ADC,

$$AD^2 + AC^2 = DC^2.$$

Como $AB = AD$, tiene que ser $BC^2 = DC^2$, y por tanto $BC = DC$. De manera que los triángulos DAC y CAB son congruentes, y el ángulo CAB, igual al CAD, debe ser recto.

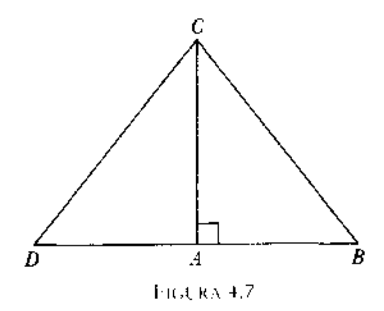


FIGURA 4.7

El material más notable del libro II es el relativo al álgebra geométrica. Ya hemos señalado que los griegos no reconocían la existencia de números irracionales, lo que les dificultaba el tratamiento numérico de longitudes, áreas, ángulos y volúmenes. En el libro II todas las cantidades están representadas geométricamente, evitando así el problema de la asignación de valores numéricos. Los números se ven sustituidos por segmentos de recta; el producto de dos números se convierte en el área del rectángulo cuyos lados tienen como

longitudes esos dos números; el producto de tres números es un volumen; la suma de dos números se traduce en la prolongación de un segmento en una longitud igual a la del otro, y la resta en recortar de un segmento la longitud del segundo; la división de un número por otro se indica por la razón entre los segmentos que los representan, de acuerdo con los principios introducidos posteriormente en los libros V y VI.

La división de un producto (un área) por un tercer número se realiza hallando un rectángulo que tenga como lado a este último y cuya área sea igual al producto dado, siendo entonces el otro lado el cociente buscado. La construcción utiliza la teoría de aplicación de áreas ya mencionada en la proposición 44 del libro I. La suma y resta de productos se reemplaza por suma y resta de rectángulos; la extracción de una raíz cuadrada, por la construcción de un cuadrado cuya área sea igual a la de un rectángulo dado (véase más adelante la proposición 14).

Las diez primeras proposiciones del libro II tratan geométricamente las proposiciones algebraicas siguientes, enunciadas con nuestro sistema notacional:

1. $a(b + c + d + ...) = ab + ac + ad + ...$;
2. $(a + b)a + (a + b)b = (a + b)^2$;
3. $(a + b)a = ab + a^2$;
4. $(a + b)^2 = a^2 + 2ab + b^2$;
5. $ab + (1/2(a + b) - b)^2 = (1/2(a + b))^2$;
6. $(2a + b)b + a^2 = (a + b)^2$.

La primera de ellas está contenida en la

Proposición 1. Si tenemos dos rectas y se divide una de ellas en un número cualquiera de partes (fig. 4.8), el rectángulo que las tiene como lados equivale a los rectángulos que tienen como lados la recta no dividida y cada una de las partes de la otra.

Las proposiciones 2 y 3 son en realidad casos particulares de la proposición 1, que Euclides trata separadamente. El conocido equivalente geométrico de 4, en palabras de Euclides, es

Proposición 4. Si se divide mediante un punto cualquiera una recta dada, el cuadrado de la recta entera es igual a los cuadrados de las partes más el doble del rectángulo que tiene a esas partes como lados.

La demostración explica los evidentes hechos geométricos contenidos en la figura 4.9.

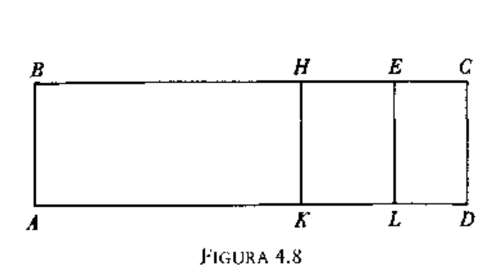


FIGURA 4.8

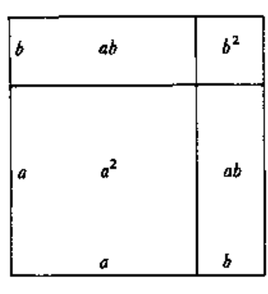


FIGURA 4.9

Proposición 11. Dividir una recta en dos partes de manera que el rectángulo que tiene como lados el total y una de las partes sea igual al cuadrado de la otra parte.

Se trata de hallar un punto H sobre el segmento AB (fig. 4.10) tal que $AB \cdot BH = AH \cdot AH$. Euclides realiza la siguiente construcción: en el cuadrado $ABDC$, toma el punto medio E del segmento AC, que une con B, y prolonga el segmento BA hasta un punto F tal que $EF = EB$; a continuación construye el cuadrado $AFGH$, y H es el punto buscado, que satisface

$$AB \cdot BH = AH \cdot AH.$$

La demostración se hace mediante áreas, utilizando teoremas anteriores, incluido el de Pitágoras, entre los que el decisivo es la proposición 6.

La importancia del teorema reside en que la longitud a del segmento AB queda dividida en longitudes x y $a - x$ tales que

$$(a - x)a = x^2$$

es decir,

$$x^2 + ax = a^2,$$

disponiendo así de un método geométrico para resolver esta última ecuación cuadrática. AB queda dividido también en media y extrema razón, ya que de $AB \cdot BH = AH \cdot AH$ se deduce que $AB : AH = AH : BH$. Otras proposiciones del libro II equivalen a la resolución de las ecuaciones cuadráticas $ax - x^2 = b^2$ y $ax + x^2 = b^2$.

Proposición 14. Construir un cuadrado equivalente a una figura rectilínea dada.

Esta última podría ser cualquier polígono; pero si es un *rectángulo* $ABEF$ (fig. 4.11), el método de Euclides equivale a lo siguiente: se prolonga AB hasta C de manera que $BC = BE$; se construye el círculo que tiene como diámetro AC y se alza en B la perpendicular DB. El cuadrado buscado es el que tiene como lado DB. Este teorema, que Euclides prueba en términos de áreas, resuelve la ecuación $x^2 = ab$, proporcionando así la raíz cuadrada de ab. En el libro IV, como veremos, se resuelven geométricamente ecuaciones cuadráticas más complicadas.

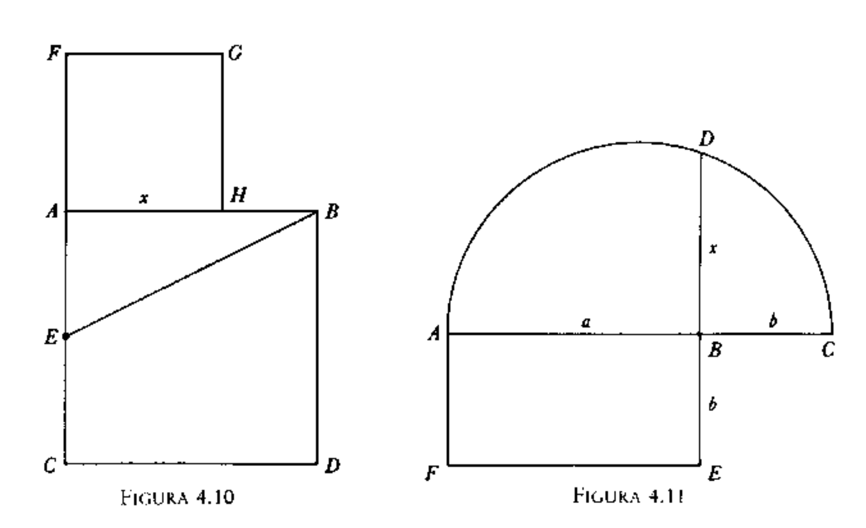


FIGURA 4.10 FIGURA 4.11

El libro III, que contiene 37 proposiciones, comienza con algunas definiciones relativas a la geometría de los círculos, y a continuación estudia las propiedades de cuerdas, tangentes, secantes, ángulos centrales e inscritos, etc. En la enseñanza media se acostumbra uno a tratar con este tipo de teoremas. El que sigue es de particular interés:

Proposición 16. La recta perpendicular en el extremo a un diámetro cae fuera del círculo, y no puede interponerse ninguna otra recta entre esa perpendicular y la circunferencia; además el ángulo del semicírculo es mayor, y el restante menor, que cualquier ángulo rectilíneo agudo.

La novedad de este teorema reside en que Euclides toma en consideración el espacio comprendido entre la tangente *TA* y el arco *ACE* (fig. 4.12); no sólo dice que es imposible trazar una recta que pase por *A* en ese espacio exterior al círculo, sino que se detiene a considerar el ángulo formado por *TA* y el arco *ACE*. Que ese ángulo, que los griegos llamaban «córneo», tuviera o no una magnitud determinada, fue un asunto controvertido. La proposición 16 afirma que es menor que cualquier ángulo rectilíneo, pero no dice que su magnitud sea nula.

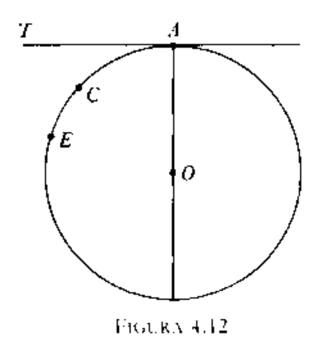


FIGURA 4.12

Proclo habla de los ángulos córneos como verdaderos ángulos. En la Edad Media y el Renacimiento [2], Cardano, Peletier, Vieta, Galileo, Wallis y otros volvieron a debatir la cuestión. Lo que hacía especialmente incómodos a los ángulos córneos para los comentaristas de Euclides es que se puedan construir círculos tangentes en *A* a *TA* de diámetro cada vez menor, cuyo ángulo córneo parece intuitivamente que debería incrementar su magnitud, lo que es negado por la proposición anterior. Por otro lado, si dos ángulos córneos cualesquiera fuesen nulos y por tanto iguales, deberían poder superponerse. Algunos comentaristas concluían de esto que los ángulos córneos no son verdaderos ángulos [3].

El libro IV trata en sus 16 proposiciones de figuras tales como triángulos, cuadrados, pentágonos y hexágonos regulares, inscritos en o circunscritos a círculos. La última proposición, que muestra cómo

[2] Bajo el nombre de «ángulo de contingencia». *(N. del T.)*

[3] Según la definición usual de ángulo entre dos curvas, el ángulo córneo es de magnitud nula.

inscribir en un círculo dado un polígono regular de 15 lados, parece haber sido usada en astronomía: hasta tiempos de Eratóstenes se creía que el ángulo de la eclíptica (el que forman el plano ecuatorial de la Tierra y el plano de su órbita en torno al Sol) era de 24°, esto es, 1/15 de 360°.

5. El libro V: La teoría de proporciones

El libro V, basado en los trabajos de Eudoxo, está considerado como el mayor logro de la geometría euclídea; su contenido y significado se han debatido más extensa e intensamente que cualquier otra porción de los *Elementos*. Se cree que los pitagóricos poseían una teoría de la proporción, esto es, de la igualdad entre dos razones, para magnitudes conmensurables: razones expresables como cociente entre dos números enteros. Aunque no conocemos los detalles de tal teoría, cabe suponer que cubría lo que veremos más tarde en el libro VII, y que se aplicaba a ciertas proposiciones sobre semejanza de triángulos. Los matemáticos que utilizaron proporciones antes de Eudoxo no poseían, en general, una fundamentación rigurosa para el tratamiento de magnitudes inconmensurables. El libro V, aun evitando la introducción de números irracionales, extiende la teoría de las proporciones a razones inconmensurables.

La noción de magnitud que presenta Euclides pretende cubrir cantidades o entidades que pueden ser conmensurables o inconmensurables entre sí: longitudes, áreas, volúmenes, ángulos, pesos, tiempo... La longitud y el área han aparecido ya, por ejemplo, en el libro II. Pero hasta ahora no ha tenido ocasión Euclides de tratar con otros tipos de magnitudes ni tampoco con sus razones mutuas o proporciones, por lo que sólo ahora introduce el concepto general de magnitud, poniendo el énfasis en las proporciones para cualquier tipo de magnitudes.

Pese a la importancia que las definiciones tienen en este libro, no hay en él una definición de magnitud como tal. Euclides comienza con la

Definición 1. Una magnitud es parte de otra mayor cuando la mide.

Parte significa aquí submúltiplo, como 2 lo es de 6, mientras que 4 no es submúltiplo de 6.

Definición 2. Lo mayor es múltiplo de lo menor cuando es medido por lo menor.

Múltiplo significa por tanto múltiplo entero.

Definición 3. Una razón es una relación entre dos magnitudes del mismo tipo con respecto a su tamaño.

Es difícil separar la significación de esta tercera definición de la que viene a continuación:

Definición 4. Se dice que hay razón entre dos magnitudes cuando se puede multiplicar cada una de ellas de manera que exceda a la otra.

Lo que significa que hay razón entre a y b si algún múltiplo entero (incluyendo 1) de a es mayor que b y algún múltiplo entero de b (incluyendo 1) es mayor que a. Esta definición excluye un concepto que apareció más tarde, el de una cantidad infinitamente pequeña y no nula, a la que se llamó infinitésimo; no cabe razón entre dos magnitudes si una de ellas es tan pequeña que ninguno de sus múltiplos enteros excede a la otra. También excluye magnitudes infinitamente grandes, a las que no superaría ningún múltiplo entero de la cantidad menor. La definición clave es la siguiente:

Definición 5. Se dice que ciertas magnitudes están en la misma razón, la primera con la segunda y la tercera con la cuarta, cuando al tomar cualquier equimúltiplo de la primera y la tercera, y cualquier equimúltiplo de la segunda y la cuarta, el múltiplo de la primera es mayor, igual o menor que el de la segunda según que el de la tercera sea mayor, igual o menor que el de la cuarta.

La definición establece que

$$\frac{a}{b} = \frac{c}{d}$$

si cuando multiplicamos a y c por *cualquier* número entero m, y b y d por *cualquier* número entero n, sean cuales fueren tales m y n,

$$ma < nb \quad \text{implica} \quad mc < nd,$$
$$ma = nb \quad \text{implica} \quad mc = nd,$$

y

$$ma > nb \quad \text{implica} \quad mc > nd.$$

Para comprender su alcance, utilicemos números modernos: para contrastar si

$$\frac{\sqrt{2}}{1} = \frac{\sqrt{6}}{\sqrt{3}}$$

deberíamos, al menos en teoría, probar que para *cualesquier* números enteros m y n,

$$m\sqrt{2} < n \cdot 1 \quad \text{implica} \quad m\sqrt{6} < n\sqrt{3}$$

y

$$m\sqrt{2} = n \cdot 1 \quad \text{implica} \quad m\sqrt{6} = n\sqrt{3}$$

y

$$m\sqrt{2} > n \cdot 1 \quad \text{implica} \quad m\sqrt{6} > n\sqrt{3}.$$

En este ejemplo, claro está, la igualdad $m\sqrt{2} = n \cdot 1$ no es posible, ya que m y n son números enteros mientras que $\sqrt{2}$ es irracional, pero esto sólo significa que la igualdad $m\sqrt{6} = n\sqrt{3}$ no tiene por qué darse; la definición establece únicamente que si alguna de las tres posibilidades de la izquierda es cierta, debe serlo también el correspondiente aserto de la derecha. Una formulación equivalente de la definición 5 sería que los enteros m y n para los que $ma < nb$ son los mismos que los enteros m' y n' para los que $m'c < n'd$.

Sería conveniente indicar inmediatamente qué partido saca Euclides de esas definiciones. Cuando se quiere probar que si $a/b = c/d$ entonces $(a + b)/b = (c + d)/d$, se consideran las razones y la proporción como números, incluso si las razones son inconmensurables, y se utiliza el álgebra para obtener el resultado; sabemos que las leyes algebraicas permiten operar con irracionales. Pero Euclides no puede hacerlo, y no lo hace; los griegos no poseían justificación suficiente para operar con razones de magnitudes inconmensurables. Así pues, Euclides prueba ese teorema usando las definiciones que ha dado, en particular la 5.ª. De hecho, está sentando las bases para un álgebra de magnitudes.

Definición 6. Las magnitudes que tienen la misma razón se llaman proporcionales.

Definición 7. Si entre los múltiplos de unas magnitudes el de la primera excede al de la segunda pero el de la tercera no excede al de la cuarta, se dice que la razón entre la primera y la segunda es mayor que la razón entre la tercera y la cuarta.

Esta definición establece que si para algunos m y n, $ma > nb$ pero mc no es mayor que nd, entonces $a/b > c/d$. Así, dada una razón entre inconmensurables a/b, se la puede situar entre otras mayores y menores que ella.

Definición 8. Una proporción tiene al menos tres términos.

En ese caso $a/b = b/c$.

Definición 9. Cuando tres magnitudes son proporcionales, se dice que la razón entre la primera y la tercera duplica la razón entre la primera y la segunda.

De modo que si $A/B = B/C$, razón entre A y C duplica la razón entre A y B, es decir, $A/C = A^2/B^2$, ya que $A = B^2/C$ y $A/C = B^2/C^2 = A^2/B^2$.

Definición 10. Cuando cuatro magnitudes son continuamente proporcionales, se dice que la razón entre la primera y la cuarta triplica la razón entre la primera y la segunda, y así sucesivamente, sea cual fuere la proporción.

O sea, que si $A/B = B/C = C/D$, la razón entre A y D triplica la razón entre A y B, es decir, $A/D = A^3/B^3$, ya que $A = B^2/C$ y $A/D = B^2/CD = (B^2/C^2)(C/D) = A^3/B^3$.

Las definiciones 11 a 18 atañen a magnitudes correspondientes, alternancia, inversión, composición, separación, conversión, etc., refiriéndose a la formación de $(a + b)/b$, $(a - b)/b$ y otras razones a partir de a/b.

El libro V prosigue con la demostración de veinticinco teoremas sobre magnitudes y razones entre magnitudes. Las pruebas son verbales y sólo dependen de las definiciones precedentes y de las nociones comunes o axiomas, tales como que al restar cosas iguales de cosas iguales se obtienen cosas iguales; no usa los postulados. Euclides emplea segmentos como ejemplos de magnitudes para ayudar al lector a comprender el significado de los teoremas y sus pruebas, pero aquéllos se aplican a toda clase de magnitudes.

Reproduciremos algunas de las proposiciones del libro V en lenguaje algebraico moderno, utilizando las letras m, n y p para los

enteros y a, b, c para las magnitudes. No obstante, para hacerse idea del lenguaje de Euclides, veamos su primera proposición:

Proposición 1. Dado cualquier número de magnitudes, sean cuales fueren, equimúltiplos de otras magnitudes en igual número, cualesquiera que fueren las veces que una de ellas sea múltiplo de alguna, ese múltiplo será de todas.

Lo que significa, en lenguaje algebraico, que $ma + mb + mc + \ldots = m(a + b + c + \ldots)$.

Proposición 4. Si $a/b = c/d$, entonces $ma/nb = mc/nd$.

Proposición 11. Si $a/b = c/d$ y $c/d = e/f$, entonces $a/b = e/f$.

Obsérvese que la igualdad entre razones depende de la definición de proporción, y Euclides pone buen cuidado en probar que la igualdad es transitiva.

Proposición 12. Si $a/b = c/d = e/f$, entonces $a/b = (a + c + e)/(b + d + f)$.

Proposición 17. Si $a/b = c/d$, entonces $(a - b)/b = (c - d)/d$.

Proposición 18. Si $a/b = c/d$, entonces $(a + b)/b = (c + d)/d$.

Algunas de estas proposiciones parecen duplicar otras del libro II. Recordemos, sin embargo, que las proposiciones de este último se referían únicamente a segmentos de recta, mientras que el libro V proporciona la teoría para toda clase de magnitudes.

El libro V fue crucial para la subsiguiente historia de las matemáticas. Los griegos clásicos no admitían números irracionales e intentaron evitarlos mediante artificios geométricos, como ya hemos indicado en nuestro repaso de los libros I a IV. Sin embargo, este uso de la geometría no tenía en cuenta las razones y proporciones de magnitudes inconmensurables de cualquier tipo, y el libro V, que inició una nueva teoría general de las magnitudes, vino a colmar esa laguna proporcionando una base firme a todo lo que en la geometría griega tuviera que ver con ellas. La cuestión clave, no obstante, es si la teoría de magnitudes servía como fundamento lógico para una teoría de los números reales que incluyera, naturalmente, a los irracionales.

Está fuera de duda cómo interpretaron a Euclides las sucesivas generaciones de matemáticos, que consideraron su teoría de las magnitudes aplicable sólo a la geometría, adoptando así la actitud de que sólo la geometría era rigurosa. Cuando se reintrodujeron los

números irracionales a partir del Renacimiento, muchos matemáticos objetaron que tales números carecían de cualquier fundamento lógico.

Un examen crítico del libro V parece darles la razón. Cierto es que las definiciones y demostraciones que Euclides presenta en el libro V no hacen uso de la geometría; como ya hemos señalado, utiliza los segmentos de recta al presentar las proposiciones y pruebas únicamente con fines pedagógicos. Aun así, si Euclides hubiera ofrecido realmente con su teoría de las magnitudes una teoría de los irracionales, ésta tendría que partir de alguna de las dos interpretaciones siguientes: o bien las magnitudes mismas, o bien las razones entre magnitudes, deberían poder ser consideradas como números irracionales.

Supongamos que las magnitudes mismas pudieran ser números irracionales. Dejando aparte cualquier crítica sobre el rigor de Euclides juzgado con criterios actuales, subsisten las siguientes dificultades: Euclides nunca define qué se entiende por magnitud, ni la igualdad o equivalencia entre magnitudes; además, él opera con proporciones, y no con las magnitudes mismas: el producto de dos magnitudes a y b sólo aparece cuando se trata de longitudes, lo que le posibilita tratarlo como un área. El producto ab no podría entonces ser un número, ya que no hay un significado general para el producto en Euclides. Además, en el libro V prueba un cierto número de teoremas sobre proporciones que podrían fácilmente ser traducidos, como más arriba hicimos, en términos algebraicos. Pero para probar la proposición 18 tiene que utilizar la cuarta proporcional a tres magnitudes dadas, lo que sólo sabe hacer para segmentos de recta (libro VI, proposición 12). Así pues, no sólo su teoría general de las magnitudes es incompleta (incluso para demostraciones que él mismo expone en el libro XII), sino que lo que establece para longitudes depende de argumentaciones geométricas. Más aún, Euclides insiste en la definición 3 en que una razón puede darse solamente entre magnitudes homogéneas. Si las magnitudes fueran números esta limitación carecería evidentemente de significado. Su concepto de magnitud, tal como es usado más tarde, está ligado a esa definición y por tanto a la geometría. Otra dificultad es que no hay un sistema de números racionales al que pudiera añadirse una teoría de los irracionales. Aparecen razones entre números enteros, pero sólo como miembros de una proporción, e incluso esas razones no son consideradas como fracciones. Por último, no existe el producto de a/b y c/d

incluso cuando las cuatro cantidades son números enteros, dejando aparte las magnitudes.

Intentemos ahora interpretar las razones entre magnitudes de Euclides como números, de modo que las razones entre inconmensurables serían los números irracionales y las razones entre conmensurables los racionales. Debería ser entonces posible al menos sumarlas y multiplicarlas. Pero en ningún momento apunta Euclides qué podría significar $(a/b) + (c/d)$ cuando a, b, c y d son magnitudes. Sus razones aparecen únicamente como elementos de una proporción, y no tienen significado general. Finalmente, como ya hemos dicho, Euclides no posee el concepto de número racional sobre el que poder construir una teoría de los irracionales.

Así pues, el curso que siguió la historia de las matemáticas hasta el siglo XIX, tratando las cantidades continuas únicamente sobre una base geométrica, era obligado. En lo que atañe a los *Elementos* de Euclides al menos, no había en ellos una fundamentación para los números irracionales.

6. El libro VI: Figuras semejantes

El libro VI, que trata de las figuras semejantes y utiliza la teoría de las proporciones del libro V, comienza con algunas definiciones. Copiaremos unas pocas [4]:

Definición 1. Figuras rectilíneas semejantes son las que tienen los correspondientes ángulos iguales, y proporcionales los lados que forman esos ángulos.

Definición 3. Una recta está dividida en extrema y media razón cuando el total es a la parte mayor como ésta a la menor.

Definición 4. La altura de cualquier figura es la perpendicular trazada desde el vértice a la base.

Esta definición es bastante imprecisa, pero Euclides la usa.

En las demostraciones de los teoremas de este libro, tal como Euclides emplea su teoría de las proporciones, no se ve obligado a

[4] Sólo falta la definición 2: Figuras inversamente proporcionales son las que tienen sus lados inversamente proporcionales a los ángulos opuestos iguales, según la traducción del griego de F. Vera, ed. Aguilar, 1970. *(N. del T.)*

tratar separadamente los casos conmensurable e inconmensurable; esta separación fue introducida por Legendre, que utilizaba una definición algebraica de proporción limitada a cantidades conmensurables, y tenía así que tratar los casos inconmensurables con otra argumentación como la *reductio ad absurdum*.

Reproduciremos aquí sólo algunos de los treinta y tres teoremas. Encontraremos de nuevo algunos resultados básicos de álgebra moderna expuestos en lenguaje geométrico.

Proposición 1. Los triángulos y paralelogramos (es decir, sus áreas) que están bajo la misma altura (que tienen la misma altitud) son entre sí como sus bases.

Euclides usa aquí una proporción entre cuatro magnitudes, dos de las cuales son áreas.

Proposición 4. En los triángulos equiángulos, los lados opuestos a los ángulos iguales son proporcionales, y también lo son los lados correspondientes que forman los ángulos iguales.

Proposición 5. Si dos triángulos tienen sus lados proporcionales, serán equiángulos y tendrán iguales los ángulos formados por los correspondientes lados.

Proposición 12. Hallar la cuarta proporcional a tres rectas dadas.

Proposición 13. Hallar la media proporcional a dos rectas dadas.

El método empleado es el corriente (fig. 4.13). Desde un punto de vista algebraico significa que, dados a y b, se puede hallar $\sqrt{ab}$.

Proposición 19. (Las áreas de) los triángulos semejantes son entre sí como la razón duplicada entre sus correspondientes lados.

Actualmente se expresa este teorema diciendo que la razón entre las áreas de triángulos semejantes es el cuadrado de la razón entre los correspondientes lados.

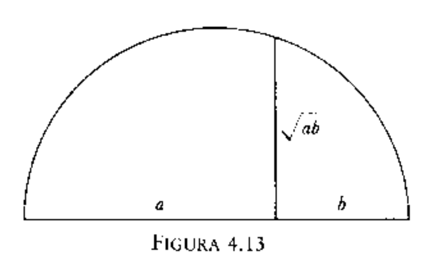


FIGURA 4.13

Proposición 27. De todos los paralelogramos aplicados a una misma recta (construidas sobre parte de esa recta) y deficientes (del construido sobre la recta entera) en paralelogramos semejantes al (paralelogramo dado) construido sobre la mitad de esa recta y similarmente dispuestos, el (de) mayor (área) es el que se aplica sobre la mitad de la recta y es semejante a su defecto.

El significado de esta proposición es el siguiente: Partiendo de un paralelogramo dado *AD* construido sobre *AC*, que es la mitad de un segmento dado *AB*, consideremos paralelogramos *AF* sobre *AK* (fig. 4.14), tales que su defecto, el paralelogramo *FB*, sea semejante a *AD*. El teorema de Euclides establece que de todos ellos el que tiene mayor área es el construido sobre *AC*.

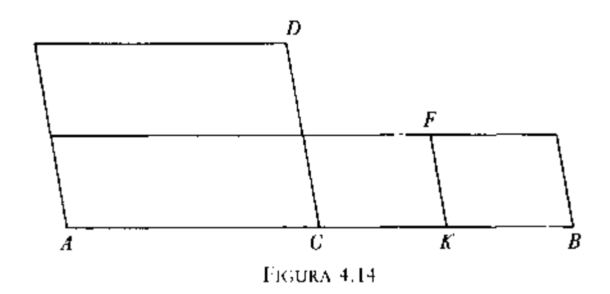


FIGURA 4.14

Esta proposición tiene un significado algebraico de gran importancia: supongamos que el paralelogramo dado *AD* sea un rectángulo (fig. 4.15) y que la razón entre sus lados es *c*/*b*, siendo *b* la longitud de *AC*; consideremos cualquier otro rectángulo *AF* que cumpla la condición de que su defecto, el rectángulo *FB*, es semejante a *AD*. Si

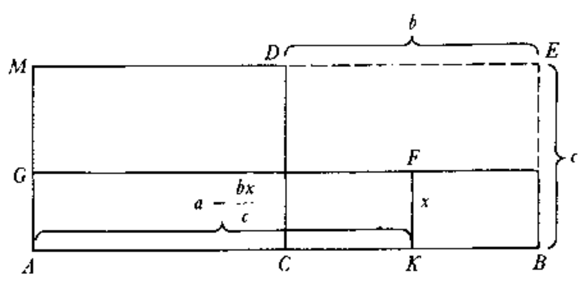


FIGURA 4.15

denotamos por x la longitud de FK, la de KB es bx/c, y si a es la longitud de AB, la de AK es $a - (bx/c)$, luego el área S de AF es

$$S = x\left(a - \frac{bx}{c}\right). \qquad (1)$$

La proposición 27 afirma que el máximo valor de S se alcanza cuando AF es AD. Como la longitud de AC es $a/2$ y la de CD es $ac/2b$, se tiene

$$S \leq \frac{a^2c}{4b}.$$

Por otro lado, para que la ecuación (1), considerada como ecuación cuadrática en x, tenga alguna raíz real, su discriminante debe ser mayor o igual que 0, esto es,

$$a^2 - 4\frac{b}{c}S \geq 0$$

o bien

$$S \leq \frac{a^2c}{4b}.$$

Así pues, la proposición no sólo nos dice cuál es el mayor valor posible de S, sino que para cada posible valor existe un x que satisface (1), y proporciona geométricamente un lado, KF, del rectángulo AF, cuya longitud es x. Este resultado se aplicará en la proposición siguiente.

Pero antes de tomarla en consideración señalemos un caso particular interesante de la proposición 27. Supongamos que el rectángulo dado AD (fig. 4.15) sea un cuadrado. En tal caso, de todos los rectángulos sobre AB cuya deficiencia sea un cuadrado, el mayor es el cuadrado sobre AC. Pero el área del rectángulo AF es $AK \cdot KF$ y como $KF = KB$, el perímetro de ese rectángulo es igual al del cuadrado DB o al del cuadrado AD, cuya área es mayor que la de AF. Así pues, de todos los rectángulos con igual perímetro el de mayor área es el cuadrado.

Proposición 28. Aplicar a una recta dada (con parte de ella como lado) un paralelogramo equivalente a una figura rectilínea dada (S) y deficiente (del paralelogramo sobre la recta entera) en un paralelogramo semejante a uno dado (D). Así (por la proposición 27), la figura rectilínea dada (S) no debe ser mayor que el paralelogramo construido sobre la mitad de la recta y semejante a su defecto.

Este teorema equivale geométricamente a la resolución de la ecuación cuadrática $ax - (b/c)x^2 = S$, donde el área S de la figura rectilínea dada está sometida, para que exista alguna solución real, a la condición $S \leq (a^2c)/(4b)$. Para comprobarlo, supongamos (porque nos conviene) que los paralelogramos son rectángulos (fig. 4.16) y sean S la figura rectilínea dada, D el otro rectángulo dado, con lados c y b, a la longitud de AB, y x la altura del rectángulo buscado. Euclides construye un rectángulo $AKFG$ de área igual a la de S tal que su defecto D' es semejante a D. Pero $AKFG = ABHG - D'$, y como D' es semejante a D su área es bx^2/c, de manera que

$$S = ax - \frac{b}{c}x^2, \tag{2}$$

y la construcción de $AKFG$ equivale a encontrar AK y x tales que x satisface la ecuación (2).

Proposición 29. Aplicar a una recta dada un paralelogramo equivalente a una figura rectilínea dada (S) y excedente en un paralelogramo semejante a uno dado (D).

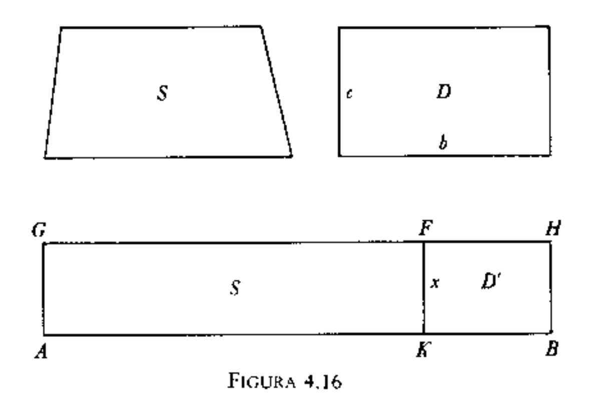


FIGURA 4.16

En términos algebraicos, este teorema resuelve

$$ax + \frac{b}{c}x^2 = S,$$

dados a, b, c y S, que ahora no está acotado porque para cualquier S positivo la ecuación tiene solución real. En lenguaje actual, Euclides muestra en las proposiciones 28 y 29 cómo resolver cualquier ecuación cuadrática en la que una o las dos raíces son positivas. Su construcción proporciona las raíces como longitudes.

Los paralelogramos construidos en las proposiciones 28 y 29 tienen un lado menor o mayor, respectivamente, que el segmento dado AB, recibiendo en griego los nombres de *Elleipsis* e *Hyperbolè*. El paralelogramo de área determinada construido sobre el segmento completo como base en la proposición 44 del libro I fue llamado *Parabolè*. Esos términos se trasladaron a las secciones cónicas por una razón que se nos hará obvia cuando estudiemos los trabajos de Apolonio.

Proposición 31. En los triángulos rectángulos, la figura construida sobre el lado opuesto al ángulo recto es equivalente a las semejantes y similarmente dispuestas sobre los lados que forman el ángulo recto.

Se trata de una generalización del teorema de Pitágoras.

7. Los libros VII, VIII y IX: La teoría de números

Los libros VII, VIII y IX tratan de la teoría de números, esto es, de las propiedades de los números enteros y de las razones entre números enteros. Son los tres únicos libros de los *Elementos* que tratan de aritmética como tal. En ellos Euclides representa los números como segmentos de recta y el producto de dos números como un rectángulo, pero sus argumentaciones no dependen de la geometría. Los asertos y pruebas son verbales, frente a la forma simbólica actual.

Muchas de las definiciones y teoremas, en particular los referidos a proporciones, repiten lo expuesto en el libro V, lo que ha llevado a los historiadores a preguntarse por qué Euclides vuelve a probar de nuevo proposiciones sobre números en lugar de aprovechar las ya probadas en el libro V.

Las respuestas son variadas. Aristóteles había incluido a los núme-

ros entre las magnitudes, pero también había enfatizado la oposición entre lo discreto y lo continuo, y no sabemos si Euclides se vio influido por las opiniones de Aristóteles sobre este tema. Tampoco se puede decidir, a partir de las vagas definiciones del libro V, si pretendía que su concepto de magnitud incluyera a los números enteros. Juzgando por el hecho de que trató a los números separadamente, parecería deducirse que éstos no figuran entre las magnitudes. Otra explicación de este estudio separado es que ya existía antes de Eudoxo una teoría de los números y las razones entre conmensurables y que Euclides respetó la tradición presentando lo que eran dos desarrollos independientes, la teoría pre-eudoxiana, en gran parte pitagórica, y la teoría eudoxiana. Pudo también creer que, dado que la teoría de números puede construirse sobre fundamentos más simples que la de las magnitudes, era sensato hacerlo. También en las contribuciones recientes a las matemáticas encuentra uno enfoques alternativos, entre los que alguno puede ser más simple que otros. Aunque Euclides separa número y magnitud, expone algunos teoremas que los relacionan. Por ejemplo, la proposición 5 del libro X establece que la razón entre dos magnitudes conmensurables es la misma que la existente entre dos números.

En estos tres libros, como en otros, Euclides da por supuestos hechos que no enuncia explícitamente; por ejemplo, que si A divide (exactamente) a B y B divide a C, entonces A divide a C; que si A divide a B y a C, también divide a $B + C$ y a $B - C$, etc.

El libro VII comienza con algunas definiciones:

Definición 3. Un número es parte de otro mayor cuando lo mide (cuando lo divide exactamente).

Definición 5. Un número es múltiplo de otro menor cuando es medido por éste.

Definición 11. Un número es primo cuando solamente lo mide la unidad.

Definición 12. Números primos entre sí son los que tienen como medida común únicamente la unidad.

Definición 13. Un número es compuesto cuando es medido por algún número (distinto de 1).

Definición 16. Cuando se multiplican dos números, el número así

obtenido se llama plano, y sus lados son los números que se han multiplicado.

Definición 17. Cuando se multiplican tres números, el número así obtenido se llama sólido, y sus lados son los números que se han multiplicado.

Definición 20. Cuatro números son proporcionales cuando el primero es el mismo múltiplo, la misma parte, o las mismas partes del segundo que el tercero del cuarto.

Definición 22. Un número es perfecto cuando es igual a (la suma de) sus propias partes.

Las proposiciones 1 y 2 exponen el proceso mediante el que se obtiene la mayor medida (divisor) común de dos números. Euclides lo describe diciendo que si A y B son los números y $B < A$, debe restarse B de A el número de veces necesario para obtener un número C menor que B. A continuación, restar C de B tantas veces como sea preciso hasta obtener un número menor que C, y así sucesivamente. Si A y B son primos entre sí se llega a 1 como último resto, y 1 es el máximo común divisor. Si A y B no son primos entre sí se llega en alguna etapa a una división exacta, y el último divisor será el mayor común. Este proceso se sigue llamando todavía algoritmo de Euclides.

Vienen a continuación teoremas simples sobre números. Por ejemplo, si $a = b/n$ y $c = d/n$, entonces $a \pm c = (b \pm d)/n$. Algunos de ellos no son sino los teoremas sobre proporciones anteriormente probados para magnitudes, y que ahora se prueban para números. Así, si $a/b = c/d$, $(a - c)/(b - d) = a/b$. En la definición 15 quedaba establecido que $a \cdot b$ es el resultado de sumar b consigo mismo a veces, y Euclides prueba ahora que $a \cdot b = b \cdot a$.

Proposición 30. Si un número primo mide al producto de dos números, debe medir al menos a uno de ellos.

Se trata de un resultado fundamental en la teoría moderna de números, cuya expresión actual se obtiene simplemente sustituyendo las palabras mide y medir por divide y dividir.

Proposición 31. Todo número compuesto es medido por algún número primo.

La demostración de Euclides parte de que si A es compuesto, por definición tiene algún divisor B; si B no es primo, es compuesto, y tiene algún divisor C que también lo es de A, etc. Y dice: «Si

proseguimos la investigación de esta forma, se encontrará algún número primo que divide al anterior, que es un divisor de A. Puesto que, si no, habría una sucesión infinita de divisores de A, cada uno de ellos menor que el anterior, y esto es imposible para los números.» Toma así en consideración el hecho de que cualquier conjunto de números enteros positivos tiene un mínimo.

El libro VIII prosigue con la teoría de números, sin incorporar nuevas definiciones. Trata sobre todo de progresiones geométricas, que para Euclides son conjuntos de números en proporción continua, esto es, $a/b = b/c = c/d = d/e = \ldots$ Tales proporciones continuas satisfacen nuestra definición de progresión geométrica, ya que en éstas la razón entre cada término y el siguiente es constante.

El libro IX concluye la tarea sobre teoría de números. Hay en él teoremas sobre números cuadrados, cúbicos, planos y sólidos, y más teoremas sobre proporciones continuas. Son de señalar los siguientes:

Proposición 14. Si un número es el menor medido por varios números primos, no puede ser medido por otros números primos.

Lo que significa que si a es el producto de los primos p, q, ... esa descomposición de a en primos es única.

Proposición 20. Hay más números primos que cualquier multitud dada de números primos.

En otras palabras, hay infinitos números primos. La demostración de Euclides es clásica: a partir de los primos $p_1, p_2, \ldots, p_n$ se puede formar el número $p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$, que es mayor que cualquiera de esos n primos, y que si es compuesto debe tener algún divisor primo diferente de todos ellos, ya que la división por $p_1, p_2, \ldots$ ó p_n deja como resto 1.

La proposición 35 del libro IX proporciona, con una elegante prueba, la suma de los términos de una progresión geométrica. La proposición 36 es un famoso teorema sobre números perfectos: si la suma de los términos de la progresión geométrica

$$1, 2, 2^2, \ldots, 2^n$$

es primo, el producto de esa suma por el último término, esto es,

$$(1 + 2 + \ldots + 2^n)2^n \quad \text{o} \quad (2^{n+1} - 1)2^n$$

es un número perfecto. Los griegos conocían los cuatro primeros números perfectos, 6, 28, 496 y 8128, y quizá también el quinto.

8. El libro X: La clasificación de los inconmensurables

El libro X de los *Elementos* emprende la tarea de clasificar en tipos los irracionales, es decir, las magnitudes inconmensurables con una magnitud dada. Augustus De Morgan describió el contenido general de este libro así: «Euclides investigó cada posible segmento cuya longitud pueda expresarse (con álgebra moderna) en la forma $\sqrt{\sqrt{a} + \sqrt{b}}$, siendo *a* y *b* las longitudes de dos segmentos conmensurables.» Claro está que no todos los irracionales pueden representarse así, y Euclides trata sólo los que surgen en su álgebra geométrica.

La primera proposición del libro X es importante para posteriores apartados de los *Elementos.*

Proposición 1. Dadas dos magnitudes desiguales, si de la mayor se resta una magnitud mayor que su mitad, y de lo que queda otra magnitud mayor que su mitad, repitiendo este proceso quedará en algún momento una magnitud menor que la más pequeña de las dos magnitudes dadas.

Al final de la demostración Euclides afirma que el teorema se puede probar igualmente si las partes sustraídas son mitades. Al principio utiliza un axioma, no reconocido como tal por Euclides, que le posibilita sumar consigo misma un número finito de veces la menor de dos magnitudes hasta obtener una suma que exceda a la mayor. Su argumentación se apoya en la definición de razón entre dos magnitudes (definición 4 del libro V), pero esa definición no justifica el paso en cuestión, ya que si sólo puede hablar de razón entre dos magnitudes cuando cada una de ellas se puede multiplicar hasta superar a la otra, Euclides debería probar que entre esas dos magnitudes existe razón, en lugar de suponerlo implícitamente. Según Arquímedes, tal axioma (aunque bajo una forma ligeramente diferente) había sido utilizado ya por Eudoxo, que lo había establecido como lema. Arquímedes lo emplea sin prueba, tomándolo de hecho por un axioma, que hoy recibe el nombre de ambos: Arquímedes-Eudoxo.

Hay 115 proposiciones en este libro X, aunque en algunas ediciones aparecen unas proposiciones 116 y 117, la última de las cuales

establece la irracionalidad de $\sqrt{2}$, con la prueba ya descrita en el capítulo 3.

9. Los libros XI, XII y XIII: Geometría de sólidos y método de exhausción

El libro XI inicia el tratamiento de los volúmenes o sólidos, aunque todavía aparecerán algunos teoremas de geometría plana. He aquí algunas de sus definiciones:

Definición 1. Un sólido es lo que tiene longitud, anchura y profundidad.

Definición 2. Los bordes de un sólido son superficies.

Definición 3. Una recta forma ángulo recto con un plano cuando lo forma con todas las rectas que la cortan y están en el plano.

Definición 4. Un plano forma ángulo recto con otro plano cuando las perpendiculares en uno de los planos a la intersección de ambos forman ángulos rectos con el otro plano.

Definición 6. La inclinación de un plano con respecto a otro es el ángulo agudo formado por las perpendiculares a la intersección común, en el mismo punto, en cada uno de los dos planos.

A ese ángulo agudo nosotros lo llamamos diedro.

Hay también definiciones para planos paralelos, figuras sólidas semejantes, ángulo sólido, pirámide, prisma, esfera, cono, cilindro, cubo, octaedro, icosaedro y dodecaedro (regulares). La esfera se define por el giro de un semicírculo en torno al diámetro que lo limita; el cono por el giro de un triángulo rectángulo en torno a uno de los lados del ángulo recto, siendo obtusángulo, rectángulo o acutángulo según que ese lado que permanece fijo en el giro sea menor, igual o mayor que el otro lado del ángulo recto; el cilindro, por el giro de un rectángulo en torno a uno de sus lados. La importancia de estas tres últimas definiciones está en que todos los sólidos considerados, excepto los poliedros regulares, se obtienen a partir del giro de una figura plana en torno a un eje.

Las definiciones son vagas, poco claras, y con frecuencia suponen teoremas no explicitados. Por ejemplo, la definición 6 da por supuesto que el ángulo es el mismo sea cual fuere el punto de la intersección de

ambos planos en que se construya. También tiende Euclides a considerar únicamente sólidos convexos, sin especificar esto en su definición de poliedro regular.

El libro sólo habla de figuras limitadas por caras planas. De los 39 teoremas que contiene, los 19 primeros se refieren a propiedades de rectas y planos, por ejemplo, acerca de rectas paralelas y perpendiculares a planos. Las demostraciones de esos teoremas en este libro no siempre son adecuadas, y muchos teoremas generales sobre poliedros sólo se prueban para ciertos casos particulares.

Proposición 20. Si un ángulo sólido está limitado por tres ángulos planos, dos cualesquiera de ellos, tomados conjuntamente de cualquier manera, son mayores que el ángulo restante.

Es decir, que de los tres ángulos planos *CAB*, *CAD* y *BAD* (fig. 4.17) la suma de dos de ellos es mayor que el tercero.

Proposición 21. Cualquier ángulo sólido está limitado por ángulos planos menores (cuya suma es menor) que cuatro ángulos rectos.

Proposición 31. Los sólidos paralelepipédicos que tienen la misma altura son entre sí como sus bases.

El libro XII contiene 18 teoremas sobre áreas y volúmenes, en particular sobre figuras curvilíneas y figuras limitadas por superficies. La idea que en él domina es la del método de exhausción, que proviene de Eudoxo. Por ejemplo, para probar que la razón entre las áreas de dos círculos es como la razón entre los cuadrados de sus diámetros, ese método aproxima ambas áreas con una precisión creciente inscribiendo en ellas polígonos regulares, y como el teorema en cuestión es válido para los polígonos, queda así probado para los

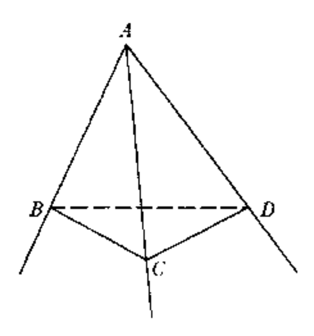


FIGURA 4.17

círculos. El término «exhausción», que proviene del hecho de que esos polígonos sucesivamente inscritos van dejando «exhausto», vacío, el círculo, no fue empleado por los griegos, sino que fue introducido en el siglo XVII. Por sí mismo, o por la vaga descripción que acabamos de dar de él, el término podría sugerir que se trata de un método aproximado, que constituye sólo una etapa hacia el concepto riguroso que se obtendría como límite. Se trata sin embargo, como vamos a ver, de un método riguroso en sí mismo, que no requiere un proceso explícito de paso al límite; su validez reside en el método indirecto de prueba, que evita el empleo de límites. De hecho, el trabajo de Euclides sobre áreas y volúmenes es más perfecto que el de Newton y Leibniz, quienes intentaron basarse en el álgebra y el sistema numérico, recurriendo a un concepto embrionario de límite.

Para una mejor comprensión del método de exhausción, consideremos con cierto detalle un ejemplo (en el próximo capítulo veremos algunos más tomados de la obra de Arquímedes). El libro XII se abre con la

Proposición 1. La razón entre los polígonos semejantes inscritos en círculos es como la razón entre los cuadrados de los diámetros de ambos círculos.

No reproduciremos la demostración porque no contiene ninguna particularidad especial. Llegamos ahora a la proposición crucial:

Proposición 2. La razón entre dos círculos es la misma que la que hay entre los cuadrados de sus diámetros.

Describiremos a continuación lo esencial de la demostración de Euclides: prueba en primer lugar que puede ir «vaciando» el círculo mediante polígonos (fig. 4.18). El área del cuadrado es mayor que la mitad del área del círculo porque aquélla es la mitad del área de un cuadrado circunscrito, que a su vez es mayor que el área del círculo. Sea ahora AB cualquiera de los lados del cuadrado inscrito, C el punto medio del arco AB, AD y BE perpendiculares a la tangente al círculo en C. $\sphericalangle 1 = \sphericalangle 2$ porque cada uno de ellos es la mitad del arco CB, de lo que se deduce que DE es paralela a AB, y $ABED$ es un rectángulo cuya área es mayor que la del segmento circular $ABFCG$. Repitiendo el proceso en cada lado del cuadrado, obtenemos un octógono regular que incluye no sólo al cuadrado sino más de la mitad de la diferencia entre el área del círculo y la del cuadrado. En cada lado del octógono podemos construir un triángulo del mismo modo que se hizo con el ACB sobre AB, obteniendo un hexadecágono regular que incluye al

octógono y más de la mitad de la diferencia entre el área del círculo y la del octógono. El proceso puede repetirse cuantas veces se desee. Euclides emplea entonces la proposición 1 del libro X para afirmar que la diferencia entre el área del círculo y la de un polígono regular con un número de lados suficientemente grande puede hacerse menor que cualquier magnitud fijada de antemano.

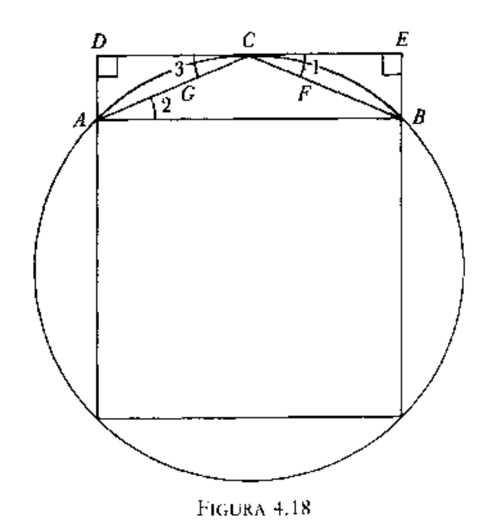


FIGURA 4.18

Sean entonces S y S' las áreas de dos círculos (fig. 4.19) y sean d y d' sus diámetros. Euclides desea probar que

$$S : S' = d^2 : d'^2. \quad (3)$$

Supongamos que no se cumple esa igualdad y que en su lugar se tiene

$$S : S'' = d^2 : d'^2, \quad (4)$$

donde S'' es algún área mayor o menor que S' (se supone aquí y en todo el libro XII la existencia de la cuarta proporcional como un área). Si $S'' < S'$, podemos construír polígonos regulares con un número cada vez mayor de lados hasta que lleguemos a uno, digamos P', tal que su área difiera de S' en menos que $S' - S''$. Ese polígono puede construirse porque ya se ha probado anteriormente que la

diferencia entre el círculo S' y los polígonos regulares inscritos en él puede hacerse menor que cualquier magnitud dada, y por tanto menor que $S' - S''$. Entonces

$$S' > P' > S''. \tag{5}$$

Inscribamos en S un polígono P semejante a P'. Por la proposición 1,

$$P : P' = d^2 : d'^2,$$

y por (4) tenemos también que

$$P : P' = S : S''$$

o bien

$$P : S = P' : S''.$$

Sin embargo, como $P < S$, resultaría

$$P' < S'',$$

en contradicción con (5).

De manera similar se puede probar que S'' no puede ser mayor que S', luego $S'' = S'$, y teniendo en cuenta (4) queda establecida la proporción (3).

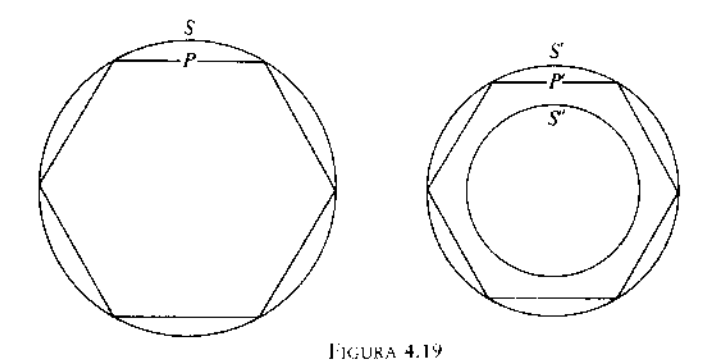


FIGURA 4.19

Este método se utiliza asimismo para probar teoremas tan críticos y difíciles como:

Proposición 5. La razón entre dos pirámides que tienen la misma altura y bases triangulares es igual a la razón entre sus bases.

Proposición 10. Cualquier cono es la tercera parte del cilindro que tiene la misma base e igual altura.

Proposición 11. La razón entre conos y cilindros de la misma altura es igual a la razón entre sus bases.

Proposición 12. La razón entre conos y cilindros semejantes es triple (razón entre los cubos) de la razón entre los diámetros de sus bases.

Proposición 18. La razón entre dos esferas es como la razón triplicada entre sus respectivos diámetros.

El libro XIII estudia propiedades de los polígonos regulares como tales e inscritos en círculos, y el problema de cómo inscribir los cinco poliedros regulares en una esfera. Prueba también que no existen más que esos cinco tipos de sólidos regulares (poliedros convexos). Este último resultado es un corolario a la proposición 18, que clausura el libro.

La prueba de que no pueden existir más que cinco tipos de sólidos regulares depende de un teorema previo, la proposición 21 del libro XI, que establece que las caras de un ángulo sólido deben sumar menos de 360°. Así, si juntamos triángulos equiláteros, podemos hacer que concurran tres en cada vértice del sólido regular para formar un tetraedro, cuatro para formar un octaedro o cinco para formar un icosaedro. Con seis triángulos equiláteros en un vértice se obtendría una suma de 360°, lo que descarta esa posibilidad. Podemos juntar tres cuadrados en cada vértice para obtener un cubo y tres pentágonos en cada vértice para formar un dodecaedro. No puede usarse ningún otro polígono regular, porque al unir tres en un punto se formaría un ángulo de 360° o más. Obsérvese que Euclides supone sólidos regulares convexos. Hay otros sólidos regulares no convexos.

Los trece libros de los *Elementos* contienen 467 proposiciones. En algunas ediciones antiguas se incluían dos libros más, que contenían otros resultados sobre sólidos regulares, aunque el libro XV es poco claro e impreciso. Ambos son, sin embargo, posteriores a Euclides. El libro XIV se debe a Hypsicles (*c.* 150 a. C.) y parte del libro XV se escribió probablemente mucho más tarde, en torno al siglo VI d. C.

10. Los méritos y defectos de los *Elementos*

Siendo los *Elementos* la primera fuente sustancial de conocimiento matemático, y habiendo sido utilizados por una generación tras otra, influyó como ningún otro libro en el derrotero de las matemáticas. Fue estudiándolos como se aprendió el concepto mismo de matemática, la noción de demostración y la ordenación lógica de los teoremas, y su contenido determinó el curso del pensamiento posterior. Por eso creemos necesario señalar las características que influyeron tan poderosamente en el futuro de la matemática.

Aunque, como ya dijimos anteriormente, la forma de presentación de las proposiciones no tiene su origen en Euclides, sí es suya la forma de presentación del conjunto de la obra: la exposición de los axiomas al comienzo, la explícita declaración de cada una de las definiciones y el ordenado encadenamiento de los teoremas, dispuestos de forma que vayan de lo más simple a lo más complejo.

Euclides también seleccionó los teoremas que consideraba de mayor importancia. Así, por ejemplo, no presenta el teorema según el cual las alturas de un triángulo se cortan en un punto. Hay teoremas en otras obras de Euclides, que discutiremos brevemente, que no consideró procedente incluir en los *Elementos.*

Aunque el requerimiento de que antes de incorporar figuras a la estructura lógica hay que demostrar su existencia es anterior a Euclides, él lo satisface con habilidad y sofisticación. De acuerdo con los postulados 1, 2 y 3, las construcciones que lleva a cabo sólo requieren el dibujo de rectas y circunferencias, con el único empleo de regla y compás. Como no pudo establecer la existencia de trisectores de un ángulo, no hay en su obra teoremas que se refieran a ellos.

Pese a algunas omisiones y errores de demostración que mencionaremos enseguida, la elección de los axiomas es notable. A partir de un pequeño número de éstos puede probar cientos de teoremas, muchos de ellos profundos. Además, esa elección fue muy inteligente, en particular en el caso del axioma de las paralelas. Sin duda sabía que cualquier axioma de ese tipo establece explícita o implícitamente lo que debe suceder en extensiones infinitas del espacio, y que cualquier pronunciamiento sobre lo que deba ser cierto en un espacio infinito es físicamente dudoso, debido a las limitaciones de la experiencia humana. Y sin embargo, también era consciente de que algún axioma de ese tenor era indispensable. Eligió por tanto una versión que establece condiciones bajo las que dos rectas se cortan en un punto a distancia

finita, y probó además todos los teoremas que pudo antes de recurrir a ese axioma.

Si bien Euclides empleó la superposición de figuras para establecer su congruencia, método que descansa en la 4.ª Noción Común, se preocupó evidentemente por la validez de tal método, al que pueden presentarse dos objeciones: en primer lugar, se utiliza el concepto de movimiento, para el que no hay una base lógica; en segundo, el método de superposición supone que una figura mantiene todas sus propiedades cuando se la mueve de un lugar a otro. Puede entonces probarse que la figura desplazada es congruente con una tercera, pero la primera figura, en su posición original, podría no serlo. Suponer que el desplazamiento de una figura no cambia sus propiedades es un requerimiento muy fuerte acerca del espacio físico. De hecho, el propósito mismo de la geometría euclídea es la comparación de figuras en diferentes posiciones. Puede constatarse esa preocupación de Euclides por la validez del método en que no lo utilizara en demostraciones que pudiera efectuar por otros medios, aunque la superposición le hubiera permitido una prueba más simple.

Aunque los matemáticos generalmente consideraron a Euclides como un modelo de rigor hasta bien entrado el siglo XIX, hay en él serios defectos que unos pocos matemáticos detectaron y combatieron. El primero es el empleo de la superposición. El segundo, la vaguedad de algunas definiciones y las imprecisiones de otras. Las definiciones iniciales de punto, línea y superficie no tienen sentido matemático preciso y, como ahora sabemos, no se les puede dar ninguno porque cualquier desarrollo matemático independiente debe incluir términos no definidos (vid. sec. 3). En cuanto a la vaguedad de muchas definiciones, basta remitirse a los comentarios que hicimos sobre el libro V, por ejemplo. Una objeción adicional a las definiciones es que varias, como la definición 17 del libro I, presuponen un axioma.

Un estudio crítico de Euclides —con la ventaja, claro está, de los conocimientos actuales— muestra que utiliza decenas de suposiciones que nunca explicita y de las que sin duda no era consciente. Algunas ya las hemos mencionado en nuestra exposición. Lo que Euclides y cientos de los mejores matemáticos de las generaciones posteriores hicieron fue emplear propiedades que las figuras sugerían como evidentes, o intuitivamente tan evidentes que no podían darse cuenta de que las estaban utilizando. En algunos casos las suposiciones

inconscientes pueden obviarse mediante demostraciones basadas en hipótesis explícitas, pero eso no siempre es posible.

Entre esas suposiciones inconscientes están las que se refieren a la continuidad de rectas y circunferencias. La demostración de la proposición 1 del libro I supone que las circunferencias tienen un punto en común. Cada una de ellas es un conjunto de puntos, y podría suceder que aunque ambas se crucen no hubiera un punto perteneciente a las dos allí donde se produce la supuesta intersección. La misma crítica puede hacerse al caso de dos rectas, que podrían cruzarse sin tener un punto común si sólo se tiene en cuenta la base lógica proporcionada por los *Elementos*.

También hay defectos en las demostraciones propuestas. Algunos son errores debidos a Euclides que pueden corregirse, aunque en ciertos casos se requeriría una nueva demostración. Otro tipo de defecto que recorre todos los *Elementos* es la afirmación de un teorema general del que sólo se prueba algún caso especial o para posiciones especiales de los datos propuestos.

Aunque hemos alabado a Euclides por la organización de conjunto del contenido de los *Elementos*, los trece libros no constituyen una unidad, sino una extensa compilación de otras obras anteriores. Por ejemplo, ya hemos señalado que los libros VII, VIII y IX repiten para los números enteros muchos de los resultados anteriormente atribuidos a las magnitudes. La primera parte del libro XIII repite resultados de los libros II y IV. Los libros X y XIII probablemente constituían una unidad, debido a Teeteto, antes de que Euclides los separara.

A pesar de estos defectos, muchos de los cuales ya fueron señalados por otros comentaristas (vid. cap. 42, sec. 1) y probablemente también por los sucesores inmediatos de Euclides, los *Elementos* tuvieron tanto éxito que desplazaron a todos los textos de geometría anteriores. En el siglo III a. C., cuando aún se disponía de otros, incluso Apolonio y Arquímedes se remitían a los *Elementos* para citar resultados anteriores a ellos.

11. Otras obras matemáticas de Euclides

Euclides escribió otras obras de matemática y física, muchas de ellas importantes para la historia de las matemáticas. Pospondremos hasta un capítulo posterior la discusión sobre sus obras de física más importantes, la *Optica* y la *Catóptrica*.

Pappus incluyó en sus *Tesoros del Análisis* los *Datos* de Euclides, describiéndolos como material geométrico suplementario relacionado con «problemas algebraicos». Los teoremas que contenía determinaban ciertas magnitudes a partir de otras dadas. Se trataba de un material de naturaleza semejante al que aparece en los *Elementos,* aunque los teoremas específicos fueran diferentes. Puede que fueran concebidos como un conjunto de ejercicios de repaso de los *Elementos.* Su contenido es íntegramente conocido.

De las obras de Euclides, a continuación de los *Elementos,* fueron las *Cónicas* las que desempeñaron un papel más relevante en la historia de las matemáticas. Según Pappus, el contenido de esta obra desaparecida, que constaba de cuatro libros, era sustancialmente el mismo que el recogido en los tres primeros libros de las *Secciones Cónicas* de Apolonio. Euclides trataba las cónicas como secciones de los tres diferentes tipos de conos (con ángulo recto, agudo y obtuso). La elipse se obtenía también como sección de cualquier cono y de un cilindro circular. Como veremos, Apolonio cambió este enfoque de las secciones cónicas.

Las *Pseudaria* de Euclides contenían demostraciones geométricas correctas y falsas, y se trataba de un libro destinado al aprendizaje de los estudiantes. La obra se ha perdido.

Sobre las divisiones (de figuras), mencionada por Proclo, trata de la subdivisión de una figura dada en otras, por ejemplo, de un triángulo en otros más pequeños o en triángulos y cuadriláteros. Existe una traducción latina, debida probablemente a Gerardo de Cremona (1114-1187), de una versión árabe incorrecta e incompleta. En 1851, Franz Woepcke encontró y tradujo otra versión árabe que parece ser correcta. Existe una traducción al inglés realizada por R. C. Archibald.

Los *Porismas* son otra obra perdida, cuyo contenido, y aun naturaleza, se desconocen en gran medida. Pappus, en su *Colección Matemática,* dice que constaba de tres libros. Se cree, basándose en los comentarios de Pappus y Proclo, que esos *Porismas* trataban esencialmente acerca de la construcción de objetos geométricos cuya existencia ya estaba asegurada. Así pues, podían considerarse como problemas intermedios entre los teoremas puros y las construcciones mediante las que se establece la existencia de alguna figura, entre los que podría ser típica la localización del centro de una circunferencia que cumpliera ciertas condiciones dadas.

La obra *Superficies-Lugares,* formada por dos libros, mencionada

por Pappus en su *Colección*, y de la que no quedan restos, trataba probablemente de lugares geométricos que son superficies.

Los *Fenómenos* de Euclides, aun siendo un texto sobre astronomía, contienen 18 proposiciones de geometría esférica y otras sobre esferas en rotación uniforme. La Tierra es tratada como una esfera. Se conservan varias versiones.

12. La obra matemática de Apolonio

El otro gran griego que pertenece al período clásico en los dos sentidos de resumir y prolongar el tipo de matemática producida en ese período es Apolonio (*c.* 262-190 a. C.). Nació en Perga, ciudad situada en el noroeste de Asia Menor, que durante su vida estaba sujeta al dominio de Pérgamo. Se trasladó a Alejandría cuando todavía era joven, y aprendió matemáticas con los sucesores de Euclides. Por lo que sabemos, permaneció en Alejandría colaborando con los grandes matemáticos que allí trabajaban. Su obra maestra es el tratado sobre las cónicas, pero también escribió sobre otros temas. Su capacidad matemática era tan extraordinaria que llegó a ser conocido en vida, y más tarde, como «el Gran Geómetra». También fue grande su reputación como astrónomo.

Las secciones cónicas, como sabemos, fueron estudiadas mucho antes de Apolonio. Concretamente, Aristeo el Viejo y Euclides habían escrito obras sobre ellas. También Arquímedes, sobre el que volveremos más adelante, presentó algunos resultados sobre este tema. Fue Apolonio, no obstante, quien lo pulió, despojándolo de irrelevancias y le dio forma sistemática. Además de sus méritos totalizadores, las *Secciones Cónicas* contienen material altamente original, y son ingeniosas, extremadamente hábiles, y están excelentemente organizadas. Se trata de una realización tan monumental que cerró prácticamente el tema para los pensadores posteriores, al menos desde el punto de vista puramente geométrico. Puede considerarse verdaderamente como la culminación de la geometría griega clásica.

Las *Secciones Cónicas* constan de ocho libros que contienen 487 proposiciones. De ellos conservamos los cuatro primeros reproducidos en manuscritos griegos de los siglos XII y XIII, y los tres siguientes en una traducción al árabe escrita en 1290. El octavo se ha perdido, aunque en el siglo XVII Halley llevó a cabo una reconstrucción basándose en las indicaciones de Pappus.

Los predecesores de Euclides, éste mismo, y Arquímedes, trataron las secciones cónicas en relación con los tres tipos de conos circulares rectos, como habían sido introducidas por el platónico Menecmo. Tanto Euclides como Arquímedes, sin embargo, sabían que la elipse también puede obtenerse como sección de los otros dos tipos de conos circulares rectos, y Arquímedes también sabía que las secciones de conos circulares *oblicuos* mediante planos que corten a todas las generatrices son elipses. Probablemente se dio cuenta de que las otras secciones cónicas pueden obtenerse a partir de conos circulares oblicuos.

Fue Apolonio, sin embargo, el primero en basar la teoría de las tres cónicas en secciones de un mismo cono circular, recto u oblicuo, y en dar cuenta de las dos ramas de la hipérbola. Se aduce como una de las razones para que Menecmo y otros predecesores de Apolonio utilizaran planos perpendiculares a una de las generatrices de los tres tipos de cono circular recto, no que no vieran que pueden obtenerse otras secciones de esos conos, sino que deseaban estudiar el problema inverso. Dadas ciertas curvas cuyas propiedades geométricas sean las de las secciones cónicas, la demostración de que esas curvas se pueden obtener como secciones de un cono es más fácil cuando el plano con el que se corta al cono es perpendicular a una generatriz.

Consideraremos en primer lugar las definiciones y propiedades básicas de las cónicas que aparecen en el libro I. Dados un círculo BC y un punto A (fig. 4.20) situado fuera del plano que contiene al círculo, una recta que pasa por A y se mueve a lo largo de la circunferencia engendra un doble cono. Al círculo se le llama base del cono. Su eje es la recta que va desde A hasta el centro del círculo (no dibujado en la figura). Si esa recta es perpendicular a la base, el cono es circular recto; si no, es escaleno u oblicuo. Consideremos la sección del cono por un plano que corte al plano de la base en una recta DE. Sea BC el diámetro del círculo base que es perpendicular a DE. Entonces ABC es un triángulo que contiene en su interior al eje del cono, y se le llama triángulo axial. Si ese triángulo corta a la cónica en PP' (que no tiene por qué ser un eje de la sección cónica), $PP'M$ es la recta determinada por la intersección del plano de corte con el triángulo axial [5]. Sea $Q'Q$ cualquier cuerda de la sección cónica

[5] Apolonio señala que si el cono es escaleno, PM no tiene por qué ser perpendicular a DE. La perpendicularidad sólo se cumple para conos circulares rectos o cuando el plano ABC es perpendicular a la base de un cono escaleno.

paralela a DE, que no tiene por qué ser perpendicular a PP'. Apolonio prueba entonces que PP' corta en el punto medio a $Q'Q$, de manera que VQ es la mitad de $Q'Q$.

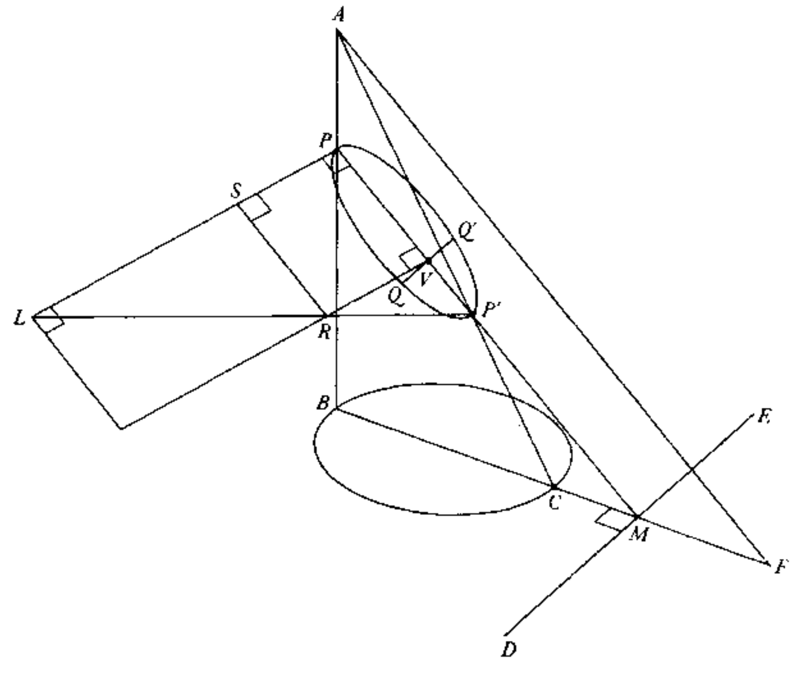


FIGURA 4.20

Dibujemos ahora la recta AF paralela a PM, hasta encontrar a BM en, digamos, F. A continuación dibujemos la recta PL perpendicular a PM en el plano de la sección. Para la elipse y la hipérbola se elige L de manera que se satisfaga la condición

$$\frac{PL}{PP'} = \frac{BF \cdot FC}{AF^2}$$

y para la parábola de manera que se tenga

$$\frac{PL}{PA} = \frac{BC^2}{BA \cdot AC}.$$

En los casos de la elipse y la hipérbola dibujemos ahora los segmentos $P'L$ y VR paralelo a PL desde V hasta cortar a $P'L$ en R (en el caso de la hipérbola P' está en la otra rama y hay que extender $P'L$ para conseguir el punto R).

Después de algunas construcciones de menor importancia que no reproduciremos, Apolonio prueba que para la elipse y la hipérbola

$$QV^2 = PV \cdot VR. \tag{6}$$

Apolonio llama a QV «ordenada» y así el resultado (6) muestra que el cuadrado de la ordenada equivale a un rectángulo construido sobre PL, en concreto el que tiene como lados PV y VR. Además, prueba que en el caso de la elipse el complementario de ese rectángulo en el rectángulo total $PV \cdot PL$ es el rectángulo LR, que es semejante al rectángulo de lados PL y PP'. De ahí el término «elipse» (vid. sec. 6).

En el caso de la hipérbola se sigue cumpliendo (6), pero la construcción mostraría que VR es más largo que PL, de manera que el rectángulo $PV \cdot VR$ excede al rectángulo construído sobre PL, esto es, $PL \cdot PV$, en un rectángulo LR que es semejante al rectángulo de lados PL y PP'. De ahí el término «hipérbola». En el caso de la parábola, Apolonio muestra que en lugar de (6) se tiene

$$QV^2 = PV \cdot PL, \tag{7}$$

de manera que el rectángulo que equivale a QV^2 es precisamente el construido sobre PL con anchura PV. De ahí el término «parábola».

Apolonio introdujo esa terminología para las cónicas en lugar de las secciones de Menecmo de los conos recto, agudo y obtuso. Cuando las palabras parábola o elipse aparecen en los trabajos de Arquímedes, como ocurre en su *Cuadratura de la Parábola* (vid. cap. 5, sec. 3), se trata de inserciones de transcriptores posteriores.

Las ecuaciones (6) y (7) son las propiedades básicas de las secciones cónicas. Una vez obtenidas, Apolonio se olvida del cono y deduce otras propiedades a partir de esas ecuaciones. De hecho, donde ahora usamos abscisa, ordenada y la ecuación de una cónica para deducir propiedades, Apolonio emplea PV, la ordenada o semicuerda QV y una igualdad geométrica, ya sea (6) o (7). Claro está que en la exposición de Apolonio no aparece nada de álgebra.

Podemos fácilmente transcribir las propiedades básicas de Apolonio en la geometría moderna con coordenadas: si denotamos por $2p$ al

segmento PL, que Apolonio llama parámetro de las ordenadas (*latus rectum* en las ediciones latinas), y por d la longitud del diámetro PP', y si x es la distancia PV medida a partir de P e y la distancia QV (lo que significa que estamos utilizando coordenadas oblicuas), se ve inmediatamente a partir de (7) que la ecuación de la parábola es

$$y^2 = 2px.$$

Para la elipse, señalemos que de la ecuación (6) que la define podemos obtener primeramente que

$$y^2 = PV \cdot VR$$

Pero $PV \cdot VR = x(2p - LS)$. También, como el rectángulo LR es semejante al determinado por PL y PP',

$$\frac{LS}{PL} = \frac{x}{d}.$$

Luego $LS = 2px/d$. Entonces

$$y^2 = x\left(2p - \frac{2px}{d}\right) = 2px - \frac{2px^2}{d}.$$

Para la hipérbola obtenemos

$$y^2 = 2px + \frac{2px^2}{d}.$$

En la construcción de Apolonio d es infinito para la parábola, y vemos así cómo ésta aparece como caso límite de la elipse o la hipérbola.

Para proseguir con el tratamiento que hace Apolonio de las cónicas necesitamos algunas definiciones de conceptos que todavía son importantes en la geometría moderna. Consideremos un conjunto de cuerdas paralelas en una elipse, digamos el conjunto de paralelas a PQ en la fig. 4.21. Apolonio prueba que los centros de esas cuerdas están en un segmento AB, al que se llama diámetro de la cónica (el segmento PP' de la figura fundamental 4.20 es un diámetro), y a

continuación, que si se dibuja una recta *DE* pasando por *C*, el punto medio de *AB*, que sea paralela a la familia original de cuerdas, esa recta corta en el punto medio a todas las cuerdas paralelas a *AB*. El segmento *DE* se llama diámetro conjugado con *AB*. En el caso de la hipérbola (fig. 4.22), las cuerdas pueden estar dentro de una de las ramas, por ej. *PQ*, y entonces el diámetro es un segmento que va de una rama a la otra, en la figura, *AB*. Las cuerdas paralelas a *AB*, por ejemplo *RH*, están entonces entre ambas ramas, y el diámetro conjugado con *AB*, digamos *DE*, definido como la media proporcional entre *AB* y el *latus rectum* de la hipérbola, no corta a la curva. En la parábola, cualquier diámetro, esto es, una recta que pase por los puntos medios de una familia de cuerdas paralelas, es siempre paralela al eje de simetría, y no hay diámetro conjugado con uno dado, ya que las cuerdas paralelas a éste son de longitud infinita. Los ejes de una elipse o hipérbola son dos diámetros conjugados perpendiculares entre sí. Para la parábola (fig. 4.23) el eje es un diámetro cuyas correspondientes cuerdas le son perpendiculares.

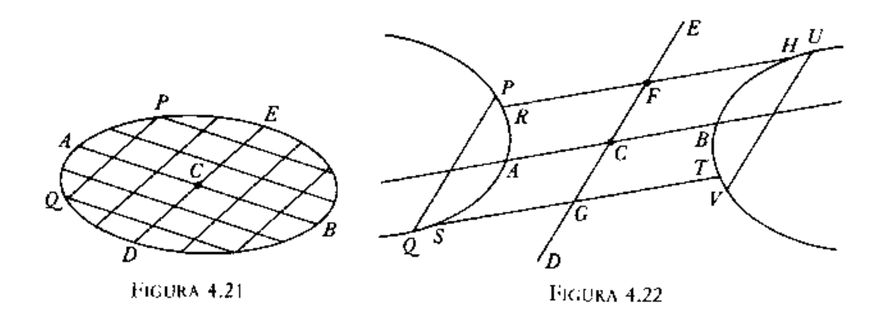


FIGURA 4.21

FIGURA 4.22

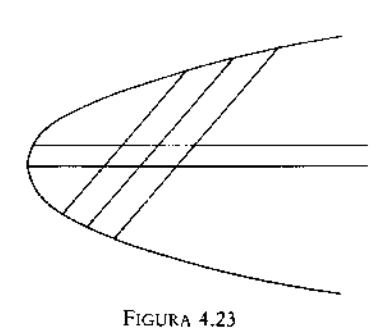


FIGURA 4.23

Después de introducir las propiedades básicas de las secciones cónicas, Apolonio presenta algunos hechos simples acerca de los diámetros conjugados. El libro I también se ocupa de las tangentes a las cónicas. Apolonio concibe una tangente como una recta que sólo tiene un punto en común con la cónica, permaneciendo cualquier otro punto fuera de ésta. Muestra entonces que si se dibuja una recta pasando por un extremo de un diámetro (punto P de la figura fundamental 4.20) que sea paralela a las cuerdas que corresponden a ese diámetro (paralelas a QQ' en esa figura), caerá fuera de la cónica, sin que pueda haber ninguna otra recta entre ella y la cónica (vid. *Elementos,* libro III, proposición 16). Por tanto, la recta mencionada es la tangente a la cónica en P.

Otro teorema sobre tangentes asegura lo siguiente: supongamos que PP' (fig. 4.24) es un diámetro de una parábola y QV es una de las cuerdas que corresponden a ese diámetro. Entonces, si se toma en él un punto T fuera de la curva y tal que $TP = PV$, donde V es el pie de la ordenada (cuerda) desde Q hasta el diámetro PP', la recta TQ será tangente a la parábola en Q. Hay teoremas análogos para la elipse y la hipérbola.

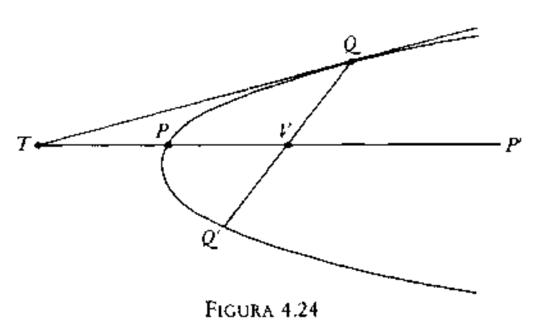


FIGURA 4.24

Apolonio prueba después que si se toma cualquier diámetro de la cónica distinto de PP' en la figura fundamental 4.20, las propiedades definitorias de la cónica, las ecuaciones (6) y (7), siguen siendo las mismas; claro está que QV se refiere entonces a la cuerda de tal diámetro. Lo que ha hecho equivale en nuestro lenguaje a una transformación de un sistema de coordenadas oblicuas en otro. En relación con esto, también prueba que a partir de cualquier diámetro y las ordenadas correspondientes se puede hacer el cambio a un diámetro (eje) cuyas ordenadas le son perpendiculares. En nuestro lenguaje, se tendría así un sistema de coordenadas rectangulares. También

muestra Apolonio cómo construir cónicas a partir de ciertos datos —por ejemplo, un diámetro, el *latus rectum*, y la inclinación de las ordenadas con respecto al diámetro—. Lo hace construyendo primeramente el cono del que la cónica deseada es una sección.

El libro II comienza con la construcción y propiedades de las asíntotas a una hipérbola. Muestra, por ejemplo, no sólo la existencia de asíntotas, sino también que la distancia entre un punto de la curva y la asíntota se hace más pequeña que cualquier longitud dada alejándose lo suficiente a lo largo de la curva. Después introduce la hipérbola conjugada con una dada, mostrando que tiene las mismas asíntotas.

Otros teoremas del libro II muestran cómo hallar un diámetro de una cónica, el centro de una cónica que lo posea, el eje de una parábola, y los ejes de una cónica con centro. Por ejemplo, si T (fig. 4.25) está fuera de una cónica dada, TQ y TQ' son tangentes en los puntos Q y Q' a la cónica, y V es el punto medio de la cuerda QQ', entonces TV es un diámetro. Otro método para encontrar un diámetro de una cónica consiste en dibujar cuerdas paralelas: la recta que une sus puntos medios es un diámetro. El punto de intersección de dos diámetros cualesquiera es el centro de la cónica (si lo tiene). El libro concluye con métodos para construir tangentes a cónicas que satisfagan ciertas condiciones dadas, como, por ejemplo, pasar por un punto dado.

El libro III comienza con teoremas sobre áreas de figuras formadas con tangentes y diámetros. Uno de los principales resultados aquí (fig. 4.26) es que si OP y OQ son tangentes a una cónica, si RS es cualquier cuerda paralela a OP y $R'S'$ cualquier cuerda paralela a OQ, y si RS y $R'S'$ se cortan en J (interna o externamente), entonces

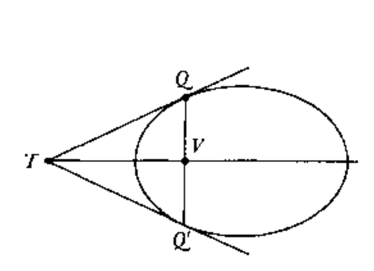


FIGURA 4.25

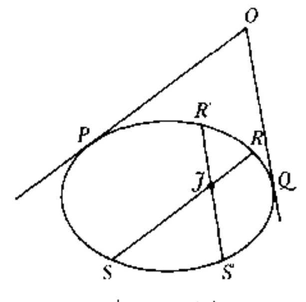


FIGURA 4.26

$$\frac{RJ \cdot JS}{R'J \cdot JS'} = \frac{OP^2}{OQ^2}\ .$$

Se trata de una generalización de un teorema bien conocido de geometría elemental, el que asegura que si dos cuerdas de un círculo se cortan, el producto de las longitudes de los segmentos producidos en una de ellas es igual al de las longitudes de los segmentos producidos en la otra, ya que en ese caso $OP^2/OQ^2 = 1$.

El libro III trata a continuación las que llamaremos relaciones armónicas entre polo y polar. Si TP y TQ son tangentes a una cónica (fig. 4.27) y si TRS es cualquier recta que corte a la cónica en R y S y a la cuerda PQ en I, se tiene

$$\frac{TR}{TS} = \frac{IR}{IS}.$$

Es decir, que T dividide a RS externamente en la misma razón en que lo hace internamente I. La recta PQ se llama polar del punto T, y se dice que T, R, I y S formar una cuaterna armónica de puntos. Por otra parte, si una recta que pase por el punto medio V del segmento PQ (fig. 4.28) corta a la cónica en R y S, y a la recta paralela a PQ que pasa por T en O, se tiene

$$\frac{OR}{OS} = \frac{VR}{VS}.$$

Esa recta que pasa por T es la polar de V, y O, R, V y S forman una cuaterna armónica de puntos.

El libro prosigue con el problema de las propiedades focales de las

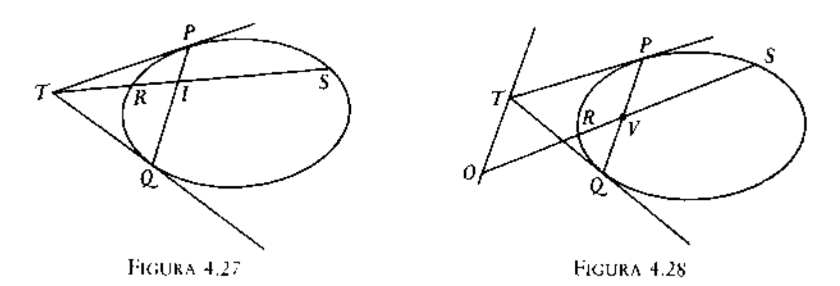


FIGURA 4.27 FIGURA 4.28

cónicas con centro; no se menciona aquí el foco de una parábola. Los focos (la palabra no es utilizada por Apolonio) se definen para la elipse y la hipérbola (fig. 4.29) como los puntos F y F' del eje (mayor) AA' tales que $AF \cdot FA' = AF' \cdot F'A' = 2p \cdot AA'/4$. Apolonio prueba para la elipse y la hipérbola que las rectas PF y PF' desde un punto P de la cónica forman ángulos iguales con la tangente en P y que la suma (para la elipse) o la diferencia (para la hipérbola) de las distancias focales PF y PF' es igual a AA'.

En esta obra no aparece el concepto de directriz, pero el hecho de que una cónica es el lugar geométrico de los puntos cuyas distancias a un punto dado (foco) y una recta dada (directriz) mantienen una razón constante ya era conocido por Euclides, y Pappus lo explicitó y demostró (vid. cap. 5, sec. 7).

Hay un problema famoso, resuelto parcialmente por Euclides, que consiste en determinar el lugar geométrico de los puntos para los que las distancias p, q, r y s a cuatro rectas dadas satisfacen la condición $pq = \alpha rs$, donde α es un número dado. Apolonio dice en su prefacio a las *Secciones Cónicas* que se puede resolver este problema con las proposiciones del libro III. Cierto es que así puede hacerse, y también en este caso Pappus sabía que ese lugar geométrico es una cónica.

El libro IV se ocupa de otras propiedades de los polos y polares. Por ejemplo, una proposición establece un método para dibujar las tangentes a una cónica desde un punto exterior T (fig. 4.30): dibuje-

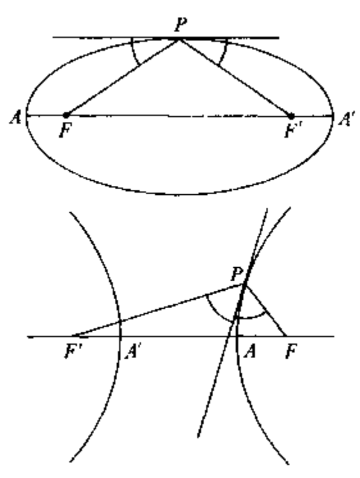


FIGURA 4.29

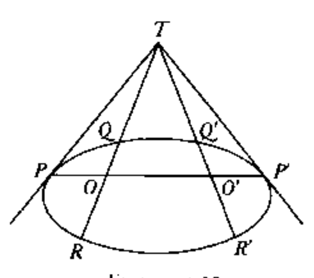


FIGURA 4.30

mos TQR y $TQ'R'$; sea O el conjugado armónico de T con respecto a Q y R, es decir, tal que $TQ : TR = OQ : OR$, y sea O' el conjugado armónico de T con respecto a Q' y R'. Dibujemos ahora OO'. Los puntos de corte P y P' son entonces los puntos de tangencia.

El resto del libro trata acerca del número de posibles intersecciones de dos cónicas en varias posiciones. Apolonio prueba que dos cónicas pueden cortarse a lo más en cuatro puntos.

El libro V es el más notable por su novedad y originalidad. Trata de las longitudes máxima y mínima de los segmentos que unen los puntos de una cónica con un punto dado. Apolonio comienza con puntos especiales sobre el eje mayor de una cónica con centro o sobre el eje de una parábola y encuentra las distancias máxima y mínima desde tales puntos a la curva. A continuación toma puntos sobre el eje menor de una elipse y hace lo mismo. Prueba también que si O es cualquier punto interior a una cónica y OP es un segmento de longitud máxima o mínima desde O hasta la cónica, la recta perpendicular en P a OP es tangente a la cónica en P; y si O' es cualquier punto sobre OP fuera de la cónica, $O'P$ es una recta mínima (segmento de longitud mínima) desde O' hasta la cónica. La perpendicular a una tangente en el punto de tangencia es lo que ahora llamamos una normal, de manera que las rectas máxima y mínima desde cualquier punto son normales. Apolonio considera a continuación propiedades de las normales a una cónica. Por ejemplo, en una parábola o una elipse, la normal en cualquier punto cortará a la curva en algún otro punto. Muestra entonces cómo se pueden construir las normales a la cónica desde un punto dado interior o exterior a la cónica.

En el transcurso de su investigación sobre los segmentos de longitud máxima y mínima (relativa) que pueden trazarse desde un punto a una cónica, Apolonio determina las posiciones de los puntos desde los que se pueden trazar dos, tres y cuatro segmentos de ese tipo. Para cada una de las cónicas determina el lugar geométrico de los puntos que separan las regiones desde las que se puede trazar uno u otro número de normales. Ese lugar mismo, que Apolonio no analiza, es lo que ahora llamamos la evoluta de la cónica, lugar geométrico de los puntos de intersección de normales a la cónica «infinitamente próximas», o envolvente de la familia de normales a la cónica. Así, desde cualquier punto dentro de la evoluta de la elipse (fig. 4.31), se pueden trazar cuatro normales a ésta, mientras que desde los puntos exteriores sólo pueden trazarse dos. (Hay puntos excepcionales.) En el caso de una parábola, la evoluta (fig. 4.32) es la curva llamada

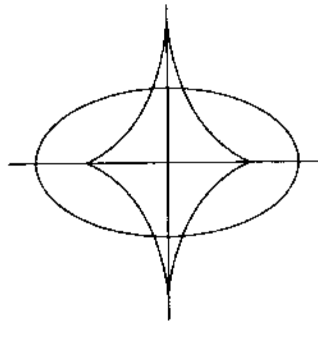


FIGURA 4.31

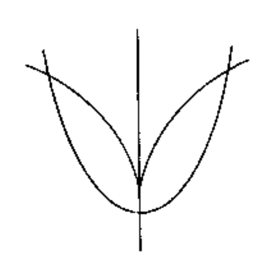


FIGURA 4.32

parábola semicúbica (estudiada por primera vez por William Neile (1637-1670)). Desde cualquier punto del plano por encima de la parábola semicúbica se pueden trazar tres normales a la parábola, y desde un punto que esté por debajo sólo una. Desde un punto de la propia parábola semicúbica se pueden trazar dos normales.

El libro VI se ocupa de cónicas y segmentos de cónicas congruentes y semejantes. Los segmentos de cónica son, como en el círculo, las regiones delimitadas por una cuerda. Apolonio muestra también cómo construir sobre un cono circular recto dado una sección cónica igual a una dada.

El libro VII no contiene proposiciones sobresalientes. Trata de propiedades de los diámetros conjugados de una cónica con centro. Apolonio compara esas propiedades con las correspondientes de los ejes. Así, si a y b son los ejes y a' y b' son dos diámetros conjugados de una elipse o hipérbola, $a + b < a' + b'$. Además, la suma de los cuadrados de dos diámetros conjugados de una elipse es igual a la suma de los cuadrados de los ejes. La proposición correspondiente para la hipérbola se obtiene reemplazando suma por diferencia. También, en una elipse o una hipérbola, el área del paralelogramo determinado por dos diámetros conjugados cualesquiera y el ángulo con el que se cortan, es igual al área del rectángulo determinado por los ejes.

El libro VIII se ha perdido. Probablemente contenía proposiciones sobre cómo determinar diámetros conjugados de una cónica (con centro) tales que ciertas funciones de sus longitudes alcancen valores dados.

Pappus menciona otras seis obras matemáticas de Apolonio. Una

de ellas, *Sobre Contactos*, cuyo contenido fue reconstruido por Vieta, contenía el famoso «problema de Apolonio»: dados tres puntos, rectas o círculos, o cualquier combinación de tres de ellos, construir una circunferencia que pase por los puntos y sea tangente a las rectas y círculos. Muchos matemáticos, incluidos Vieta y Newton, proporcionaron soluciones a este problema.

La matemática estrictamente deductiva de Euclides y Apolonio ha alentado la impresión de que los matemáticos crean razonando deductivamente. Nuestro repaso a los trescientos años de actividad anteriores a Euclides muestra, sin embargo, que las conjeturas precedieron a las pruebas y el análisis a la síntesis. De hecho, los griegos no concedían mucho mérito a las proposiciones obtenidas mediante simple deducción. A los resultados que se derivan fácilmente de un teorema los llamaron corolarios o porismas. Tales resultados, obtenidos sin mucho trabajo adicional, fueron considerados por Proclo como frutos caídos del árbol o propinas.

No hemos agotado las contribuciones del genio griego a la matemática. Lo que hemos discutido hasta ahora pertenece al período griego clásico; todavía nos espera la importante época que se extiende desde el año 300 a. C., más o menos, hasta el 600 d. C. Antes de volver la página insistiremos en que el período clásico contribuyó con algo más que sus contenidos: creó la matemática misma en el sentido en que hoy entendemos la palabra. La insistencia en la deducción como método de demostración y la preferencia por lo abstracto en oposición a lo concreto determinaron el carácter de las matemáticas, mientras que la selección del conjunto de axiomas más fructífero y aceptable, y la intuición y demostración de cientos de teoremas pusieron en marcha esta ciencia.

Bibliografía

Ball, W. W. Rouse: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, capítulos 2 y 3.

Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Coolidge, Julian L.: *A History of Geometrical Methods*, Dover (reimpresión), 1963, libro I, capítulos 2 y 3.

Heath, Thomas L.: *A Manual of Greek Mathematics*, Dover (reimpresión), 1963, capítulos 3 al 9 y 12.

—: *A History of Greek Mathematics*, Oxford University Press, 1921, vol. I, capítulos 3 al 11; vol. II, capítulo 14.
—: *The Thirteen Books of Euclid's Elements*, en 3 vols., Dover (reimpresión), 1956.
—: *Apolonius of Perga*, Barnes & Noble (reimpresión), 1961.
Neugebauer, Otto: *The Exact Sciences in Antiquity*, Princeton University Press, 1952, capítulo 6.
Proclo: *A Commentary on the First Book of Euclid's Elements*, Princeton University Press, 1970.
Sarton, George: *A History of Science*, Harvard University Press, 1952 y 1959, vol. I, capítulos 8, 10, 11, 17, 20, y capítulo 3 del vol. II.
Scott, J. F.: *A History of Mathematics*, Dover (reimpresión), 1956, vol. I, capítulo 3; vol. II, capítulo 5.
Struik, Dirk J.: *A Concise History of Mathematics*, 3.ª edición, Dover, 1967, cap. 3.
Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, capítulos 4 al 6.

Capítulo 5

EL PERIODO GRECO-ALEJANDRINO: GEOMETRIA Y TRIGONOMETRIA

Sin los conceptos, métodos y resultados hallados y desarrollados por generaciones precedentes desde la antigüedad griega, no podemos comprender ni los objetivos ni las conclusiones de las matemáticas en los últimos cincuenta años.

HERMANN WEYL

1. La fundación de Alejandría

La evolución de la matemática ha estado fuertemente ligada al curso de la historia. Las conquistas acometidas por los macedonios, un pueblo griego que vivía en la parte septentrional de las tierras de Grecia, llevó consigo la destrucción de la civilización clásica griega y puso las bases de otra civilización, esencialmente griega pero de carácter completamente diferente. Las conquistas fueron iniciadas el año 352 a. C. por Filipo II de Macedonia. Atenas fue derrotada el año 338 a. C. El año 336 a. C. Alejandro Magno, hijo de Filipo, tomó el mando y conquistó Grecia, Egipto y el Oriente Próximo, llegando por el este hasta la India y por el sur hasta las cataratas del Nilo. Construyó nuevas ciudades por todas partes, que eran a la vez fortalezas y centros de comercio. La más importante de todas, Alejandría, situada en el centro del imperio de Alejandro y con la intención de ser su capital, fue fundada en Egipto el año 332 a. C. Alejandro eligió el lugar y dibujó los planos para la construcción y la colonización de la ciudad, pero el trabajo no fue completado hasta muchos años después.

Alejandro imaginaba una cultura cosmopolita en su nuevo imperio. Debido a que, entre las demás, la única civilización importante era

la persa, Alejandró intentó deliberadamente fundir ambas culturas. El año 325 a. C., él mismo se casó con Statira, hija del príncipe persa Darío, e indujo a cien de sus generales y a diez mil de sus soldados a casarse con mujeres persas. Incorporó veinte mil persas a su ejército y los mezcló con los macedonios en las mismas falanges. Asimismo, llevó colonizadores de todas las naciones a las diferentes ciudades fundadas por él. Tras su muerte, se encontraron órdenes escritas de transportar grandes grupos de europeos a Asia y viceversa.

Alejandro murió el año 323 a. C., antes de terminar su capital y cuando estaba todavía ocupado en sus conquistas. Después de su muerte, sus generales se enfrentaron entre sí para conseguir el poder. Tras varias décadas de inestabilidad política, el imperio se descompuso en tres partes independientes. La parte europea constituyó el imperio Antigónido (del general griego Antigono); la parte asiática, el imperio Seléucida (por el general Seleuco), y Egipto, gobernado por la dinastía griega Ptolemaica se convirtió en el tercer imperio. Antigonia, Grecia y Macedonia fueron cayendo de forma gradual bajo la dominación romana y su importancia, en lo que concierne al desarrollo de la matemática, llegó a ser insignificante. La matemática desarrollada en el imperio Seléucida fue principalmente una continuación de la matemática babilónica, completamente influida por los acontecimientos que estamos considerando. Las creaciones más importantes que continuaban el período clásico griego tuvieron lugar en el imperio Ptolemaico, principalmente en Alejandría.

El hecho de que el imperio Ptolemaico se convirtiera en el heredero matemático de la Grecia clásica no fue accidental. Los reyes del imperio fueron griegos sabios y continuaron el plan de Alejandro de construir un centro cultural en Alejandría. Ptolomeo Soter, que reinó del 323 a. C. al 285 a. C., sus inmediatos sucesores, Ptolomeo II, llamado Filadelfo, que reinó del 285 a. C. al 247 a. C., y Ptolomeo Euergetes, que lo hizo del 247 a. C. al 222 a. C. estaban muy bien enterados de la importancia cultural de las grandes escuelas griegas tales como las de Pitágoras, Platón y Aristóteles. Estos gobernantes llevaron a Alejandría estudiosos de todos los centros de cultura existentes y los mantuvieron mediante ayudas estatales. Alrededor del año 290 a. C., Ptolomeo Soter construyó un centro en el cual los sabios estudiarían y enseñarían. Este edificio, dedicado a las musas, fue conocido como el Museo y albergó a poetas, filósofos, filólogos, astrónomos, geógrafos, médicos, historiadores, artistas y la mayoría de los matemáticos famosos de la civilización greco-alejandrina.

Junto al Museo, Ptolomeo construyó una biblioteca, no sólo para la conservación de documentos importantes sino también para uso de todo tipo de público. Esta famosa biblioteca llegó a tener 750.000 volúmenes a un tiempo, e incluía las bibliotecas personales de Aristóteles y de su sucesor, Teofrasto. Los libros, casualmente, eran más fáciles de obtener en Alejandría que en la Grecia clásica debido a que el papiro egipcio estaba más a mano. De hecho, Alejandría se convirtió en el centro de fabricación de libros del mundo antiguo.

Los Ptolomeos continuaron también el plan de Alejandro de fomentar una fusión entre los pueblos, por lo que griegos, persas, judíos, etíopes, árabes, romanos, hindúes y negros se desplazaron a Alejandría sin encontrar obstáculos y se confundieron libremente en la ciudad. Aristócratas, ciudadanos y esclavos convivieron entre sí y, de hecho, las distinciones de clase de la vieja civilización griega desaparecieron. La civilización de Egipto recibió la influencia de los conocimientos que llevaron los mercaderes y las expediciones especiales organizadas por los sabios para aprender más cosas acerca de otras partes del mundo. En consecuencia, los horizontes intelectuales se ensancharon. Los largos viajes por mar de los alejandrinos necesitaban un mejor conocimiento de la geografía, de los métodos de medición del tiempo y de las técnicas de navegación, mientras que la competencia comercial generó el interés por los materiales, por el rendimiento de la producción y por el perfeccionamiento de los especialistas. Artes que habían sido despreciadas en el período clásico renacieron con nuevos bríos y se crearon escuelas de perfeccionamiento. La ciencia pura siguió cultivándose, pero también hizo su aparición la ciencia aplicada.

Los aparatos mecánicos creados por los alejandrinos resultan sorprendentes incluso para criterios modernos: bombas para elevar agua desde pozos y cisternas, poleas mecánicas, cuñas, poleas marinas, sistemas de engranajes, y un cuentamillas en absoluto diferente de los que se pueden encontrar en un coche moderno se usaban de manera habitual. La fuerza del vapor se empleaba para conducir un vehículo a lo largo de las calles de la ciudad durante la procesión religiosa anual. El agua o el aire calentados mediante el fuego en vasijas ocultas en los altares del templo se utilizaban para fabricar estatuas móviles. El público observaba atónito cómo los dioses levantaban sus manos para bendecir a los fieles, dioses derramando lágrimas y estatuas lanzando bocanadas de agua. La fuerza del agua accionaba un órgano musical y trazaba figuras automáticamente en

una fuente mientras el aire comprimido se usaba para hacer funcionar un cañón. Con objeto de mejorar las mediciones astronómicas se inventaron nuevos instrumentos mecánicos, incluido un reloj de sol muy perfeccionado.

Los alejandrinos tenían un conocimiento avanzado de fenómenos tales como el sonido y la luz. Conocían la ley de la reflexión y tenían un conocimiento empírico de la ley de la refracción (cap. 7, sec. 7), conocimientos que aplicaron a la construcción de espejos y lentes. Durante este período tuvo lugar la aparición por primera vez de un trabajo de metalurgia, que llevaba consigo una carga mucho mayor de química que los pocos hechos que conocían los antiguos sabios egipcios y griegos. Los venenos fueron una especialidad. La medicina floreció, debido en parte a que la disección del cuerpo humano, prohibida en la Grecia antigua, estaba permitida ahora, y el arte de la curación alcanzó su cumbre con la obra de Galeno (129-*c.* 201), quien, no obstante, vivió principalmente en Pérgamo y Roma. La Hidrostática, la ciencia del equilibrio de los cuerpos sumergidos en fluidos fue investigada con intensidad y, naturalmente, fundamentada de manera sistemática. El mayor de sus logros científicos fue la primera teoría astronómica verdaderamente cuantitativa (cap. 7, sec. 4).

2. El carácter de la matemática greco-alejandrina

El trabajo de los sabios en el Museo estaba dividido en cuatro secciones: literatura, matemáticas, astronomía y medicina. Puesto que dos de ellas eran esencialmente matemáticas y la medicina, a través de la astrología, precisa de algunas matemáticas, vemos que éstas ocupaban un lugar preponderante en el mundo alejandrino. Las características de la matemática estuvieron muy influidas por las nuevas civilización y cultura. Pese a lo que puedan decir los matemáticos acerca de la pureza de sus temas y su indiferencia en lo que se refiere a, y una elevación respecto de, su entorno social, la nueva civilización helenística produjo una matemática de características completamente diferentes de las del período clásico.

Naturalmente, Euclides y Apolonio fueron alejandrinos; pero, como ya hemos observado, Euclides organizó el trabajo del período clásico, y Apolonio es excepcional, ya que también organizó y extendió la matemática griega clásica —pese a que en su astronomía y sus trabajos sobre los números irracionales (que presentaremos en

posteriores capítulos), estuvo influido a veces por la cultura alejandrina. Con toda seguridad, los restantes grandes matemáticos alejandrinos, Arquímedes, Eratóstenes, Hiparco, Nicomedes, Herón, Menelao, Ptolomeo, Diofanto y Pappus desplegaron el genio griego para la matemática teórica y abstracta, pero con notables diferencias. La geometría alejandrina se dedicaba principalmente a la obtención de resultados útiles para el cálculo de longitudes, áreas y volúmenes. Ciertamente, algunos de estos teoremas aparecen también en los *Elementos* de Euclides. Por ejemplo, la proposición 10 del libro XII señala que todo cono es la tercera parte del cilindro que tiene su misma base e igual altura. Por tanto, si se conoce el volumen de un cilindro se puede saber el de un cono. Sin embargo, tales teoremas son relativamente escasos en Euclides, mientras que ocupan la mayor parte de la atención de los geómetras alejandrinos. Así, mientras Euclides se contentaba con probar que la razón de las áreas de dos círculos es igual a la de los cuadrados de sus diámetros respectivos —lo que nos permite saber que el área es $A = k \cdot d^2$, pero sin un valor de k— Arquímedes obtuvo una aproximación muy exacta del número π, con lo que se podían calcular las áreas circulares.

Además, los griegos clásicos, debido a que no tomaban en consideración los números irracionales, produjeron una geometría estrictamente cualitativa. Los alejandrinos, de acuerdo con la práctica de los babilonios, no dudaron en usar los irracionales y asignar libremente números a longitudes, áreas y volúmenes. La culminación de estos trabajos fue el desarrollo de la trigonometría.

Incluso más significativo fue el hecho de que los alejandrinos resucitaron y extendieron la aritmética y el álgebra, que se convirtieron en temas de pleno derecho. Este desarrollo de la ciencia de los números era, por supuesto, imprescindible si se pretendía obtener un conocimiento cuantitativo tanto de los resultados geométricos como del uso directo del álgebra.

Los matemáticos alejandrinos tomaron parte activa en trabajos de mecánica. Calculaban centros de gravedad de cuerpos de distintas formas; trabajaban con fuerzas, planos inclinados, poleas y engranajes, y a menudo se convertían en inventores. Eran también los principales contribuidores de su época en trabajos sobre la luz, geografía matemática y astronomía.

En el período clásico la matemática abarcaba la aritmética (sólo de los números enteros), la geometría, la música y la astronomía. El panorama de la matemática sufrió una gran expansión en el período

alejandrino. Proclo, que importó material procedente de Gémino de Rodas (siglo I a. C.) cita la última división de las matemáticas (seguramente en la época de Gémino): aritmética (nuestra teoría de números), geometría, mecánica, astronomía, óptica, geodesia, canon (ciencia de la armonía musical) y logística (aritmética aplicada). De acuerdo con Proclo, Gémino dice: «La totalidad de la matemática estaba dividida en dos grandes apartados con las siguientes distinciones: una parte relativa a los conceptos intelectuales propios y otra a los conceptos materiales.» La aritmética y la geometría eran intelectuales. La parte restante era material. No obstante, esta distinción fue disminuyendo progresivamente, pero a finales del siglo I a. C. todavía era significativa. Podemos decir, en una generalización poco rigurosa, que las matemáticas del período alejandrino cortaron su relación con la filosofía y se aliaron con la ingeniería.

Trataremos en primer lugar de los trabajos alejandrinos sobre geometría y trigonometría. En el capítulo siguiente discutiremos la aritmética y el álgebra.

3. Areas y volúmenes en los trabajos de Arquímedes

No hay ninguna persona cuyos trabajos sinteticen el carácter de la edad alejandrina tan bien como Arquímedes (287-212 a. C.), el mayor matemático de la antigüedad. Hijo de un astrónomo, había nacido en Siracusa, un asentamiento griego en Sicilia. En su juventud fue a Alejandría, donde recibió su educación. Pese a que regresó a Siracusa y pasó allí el resto de su vida, estuvo en contacto con Alejandría. Era muy conocido en el mundo griego y fue muy admirado y respetado por sus contemporáneos.

Arquímedes estaba en posesión de una inteligencia sublime, una gran amplitud de intereses —tanto prácticos como teóricos— y una excelente habilidad mecánica. Sus trabajos en matemáticas incluyen el cálculo de áreas y volúmenes por el método de aproximaciones sucesivas, el cálculo del número π (en el transcurso del cual aproximó raíces cuadradas de números grandes y pequeños), y un sistema nuevo para representar números grandes en el lenguaje oral. En mecánica calculó los centros de gravedad de varias figuras planas y sólidas y dio teoremas sobre la palanca. La parte de la hidrostática que trata del equilibrio de los cuerpos que flotan en el agua fue creada por él. También tiene fama de haber sido un buen astrónomo.

Sus descubrimientos rebasaron en tal medida la técnica de su tiempo que a su alrededor surgieron un sinfín de historias y leyendas. En realidad, en la estima popular sus inventos oscurecieron sus matemáticas, pese a que puede situarse con Newton y Gauss como uno de los tres más grandes en este campo. En su juventud construyó un planetario, un mecanismo que funcionaba gracias a la potencia del agua y que reproducía los movimeintos del Sol, la Luna y los planetas. Ideó una bomba (la hélice de Arquímedes) para elevar agua desde un río; mostró cómo usar la palanca para mover grandes pesos; utilizó poleas compuestas para botar una galera para el rey Hierón de Siracusa, e inventó ingenios militares y catapultas para proteger Siracusa cuando fue atacada por los romanos. Aprovechando las propiedades focales de un espejo en forma de paraboloide, incendió las naves romanas que sitiaban Siracusa concentrando sobre ellas los rayos solares.

Seguramente la historia más famosa sobre Arquímedes sea su descubrimiento del método para determinar la falsificación de una corona de oro. El rey de Siracusa había encargado la corona. Cuando se la entregaron sospechó que en su interior no habían colocado metales nobles y la hizo llegar a Arquímedes para que encontrara algún procedimiento que permitiera determinar su contenido sin que, por supuesto, hubiera que destruir la pieza. Arquímedes se planteó el problema; un día, mientras se estaba bañando observó que su cuerpo sufría un empuje hacia arriba producido por el agua y de repente comprendió el principio que le iba a permitir dar una solución al problema. Estaba tan excitado por su descubrimiento que iba dando saltos por la calle gritando «¡Eureka!» («¡lo he encontrado!»). Había descubierto que un cuerpo sumergido en el agua sufre un empuje vertical hacia arriba con una fuerza igual al peso del agua desalojada, y mediante este principio fue capaz de determinar la composición de la corona (ver cap. 7, sec. 6).

Pese a que Arquímedes era notablemente ingenioso y un inventor de fama, Plutarco dice que estos inventos no eran más que «la diversión del geómetra». Según Plutarco, Arquímedes «estaba en posesión de un espíritu tan alto, un alma tan profunda y una riqueza tal de conocimientos científicos que, a pesar de que estos inventos le habían proporcionado la celebridad de tener más que sabiduría humana, no dejaría tras él ningún trabajo escrito sobre tales cuestiones, sino que, considerando como innobles y viles los trabajos mecánicos y todo tipo de arte que se puede usar y aprovechar directamente, centró

su mayor ambición en aquellas especulaciones cuya belleza y sutileza no añaden nada a las necesidades habituales de la vida». Sin embargo, la importancia de Plutarco como relator de historias es mucho mayor que como historiador. Arquímedes escribió libros sobre mecánica entre los que tenemos el que se titula *Sobre la flotación de los cuerpos* y otro, *Sobre el equilibrio de planos;* otros dos, *Sobre palancas* y *Sobre centros de gravedad* se han perdido. Escribió también un trabajo sobre óptica que ha desaparecido y trataba de sus descubrimientos; aunque el trabajo se ha perdido se sabe con certeza que escribió *Sobre la estructura de la esfera*, que describe un invento que muestra los movimientos del Sol, la Luna y los cinco planetas alrededor de la Tierra (fija).

La muerte de Arquímedes fue un presagio de lo que iba a suceder en todo el mundo griego. El año 216 a. C. Siracusa se alió con Cartago en la segunda guerra Púnica entre esa ciudad y Roma. Los romanos atacaron Siracusa el año 212 a. C. Mientras estaba dibujando figuras matemáticas en la arena, uno de los soldados romanos que acababan de tomar la ciudad dio el alto a Arquímedes. El caso es que Arquímedes se sintió confuso aunque se hizo el sordo ante el aviso del soldado romano. Tras esto, el soldado lo mató, a pesar de la orden del comandante romano, Marcelo, de que se le respetase la vida. Tenía entonces setenta y cinco años y estaba todavía en perfecta posesión de todas sus facultades. A modo de «compensación», los romanos construyeron una tumba muy historiada sobre la cual inscribieron un famoso teorema arquimediano.

Los escritos de Arquímedes toman la forma de pequeños tratados en vez de grandes libros. Nuestro conocimiento de estos trabajos viene de los manuscritos griegos existentes y de los manuscritos latinos traducidos del griego del siglo XIII en adelante. Alguna de las versiones latinas se hicieron a partir de manuscritos griegos asequibles a los traductores, pero no para nosotros. En 1543 Tartaglia hizo una traducción al latín de algunos trabajos de Arquímedes.

Los trabajos geométricos de Arquímedes representan el cenit de la matemática greco-alejandrina. En sus razonamientos matemáticos, Arquímedes usa teoremas de Euclides y Aristeo, así como otros resultados que él dice que son evidentes, es decir, pueden probarse fácilmente a partir de resultados conocidos. Sus demostraciones están perfectamente razonadas pero no resultan fáciles para nosotros ya que no estamos familiarizados con muchos de los métodos y resultados de los geómetras griegos.

En su trabajo *Sobre la esfera y el cilindro* Arquímedes comienza con definiciones e hipótesis. La primera hipótesis o axioma es que de entre todas las líneas (curvas) que tienen los mismos extremos la línea recta es la más corta. Otros axiomas se refieren a longitudes de curvas cóncavas y superficies. Por ejemplo, *ADB* (fig. 5.1) se supone que es menor que *ACB*. Estos axiomas conducen a Arquímedes a comparar perímetros de polígonos inscritos y circunscritos con el perímetro del círculo.

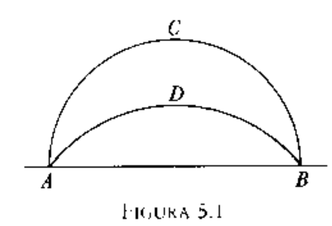


FIGURA 5.1

Después de algunas proposiciones preliminares, en el libro I prueba:

Proposición 13. La superficie de cualquier cilindro circular recto sin incluir las bases es igual a [el área de] un círculo cuya base es media proporcional entre el lado [una generatriz] y el diámetro de su base.

Esto viene seguido de varios teoremas relativos al volumen de conos. De gran interés son:

Proposición 33. La superficie de cualquier esfera es cuatro veces el [área de] uno de sus círculos máximos.
Corolario a la proposición 34. Todo cilindro cuya base es un círculo máximo de una esfera y cuya altura es igual al diámetro de la esfera es 3/2 de [el volumen de] la esfera, y su superficie junto con sus bases es 3/2 de la superficie de la esfera.

Es decir, compara el área de la superficie y el volumen de una esfera con un cilindro circunscrito a la misma. Este es el famoso teorema que, de acuerdo con los deseos de Arquímedes, se inscribió sobre su lápida.

Prueba después en las proposiciones 42 y 43 que la superficie del segmento esférico *ALMNP* es el área de un círculo cuyo radio es *AL* (fig. 5.2). El segmento puede ser menor o mayor que una semiesfera.

El teorema del área de la superficie y el volumen se prueba por el

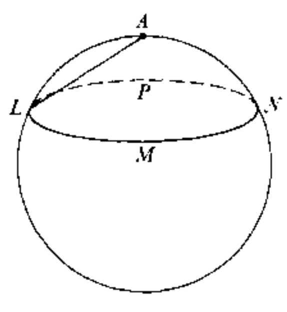


FIGURA 5.2

método de las aproximaciones sucesivas. Arquímedes utiliza figuras rectilíneas inscritas y circunscritas para «agotar» el área o el volumen y entonces, igual que Euclides, usa el método indirecto de demostración para completar el argumento.

Algunos teoremas del segundo libro de *Sobre la Esfera y el Cilindro* que se refieren sobre todo a segmentos esféricos son significativos, pues contienen una nueva álgebra geométrica. Por ejemplo, enuncia:

Proposición 4. Cortar una esfera con un plano de manera que los volúmenes de los segmentos obtenidos estén en una razón dada.

Este problema lleva algebraicamente a la resolución de la ecuación cúbica:

$$(a - x) : c = b^2 : x^2$$

y Arquímedes la resuelve geométricamente hallando la intersección de una parábola y una hipérbola rectangular.

El trabajo *Sobre Conoides y Esferoides* estudia propiedades de figuras de revolución generadas por cónicas. El conoide de ángulo recto de Arquímedes es un paraboloide de revolución. (En tiempos de Arquímedes se consideraba todavía la parábola como una sección de un cono de ángulo recto.) El conoide de ángulo obtuso es una rama de un hiperboloide de revolución. Los esferoides de Arquímedes son lo que llamamos esferoides achatado y oblongo, que son figuras de revolución generadas por elipses. El objetivo principal del trabajo es la determinación de volúmenes de segmentos obtenidos al cortar cuerpos tridimensionales con planos. El libro contiene también algún

trabajo de Arquímedes acerca de las secciones cónicas, ya citado al hablar de Apolonio. Como en otros trabajos, presupone teoremas que considera probados con facilidad o que pueden probarse con procedimientos usados con anterioridad. Varias de las demostraciones utilizan el método de las aproximaciones sucesivas. Algunos ejemplos de los contenidos pueden hallarse en las siguientes proposiciones:

Proposición 5. Si AA' y BB' son los ejes mayor y menor de una elipse y si d es el diámetro de cualquier círculo, el área de la elipse es al área del círculo como $AA' \cdot BB'$ es a d^2.

El teorema dice que si $2a$ es el eje mayor y $2b$, el eje menor y s y s' son las áreas de la elipse y el círculo respectivamente, entonces $s/s' = 4ab/d^2$, ya que $s' = (\pi/4)d^2$, $s = \pi ab$.

Proposición 7. Dadas una elipse de centro C y una línea CO perpendicular al plano de la elipse es posible encontrar un cono circular de vértice O de manera que la elipse es una sección del mismo.

Claramente, Arquímedes da por cierto que algunas, al menos, de las distintas secciones cónicas pueden obtenerse de un mismo cono, hecho utilizado ya por Apolonio.

Proposición 11. Si un paraboloide de revolución se corta por un plano que contiene al eje [de revolución], o es paralelo al mismo, la sección será una parábola igual a la parábola original que genera el paraboloide... Si se corta el paraboloide por un plano perpendicular a su eje la sección será un círculo cuyo centro está en el eje.

Hay resultados análogos para el hiperboloide y el esferoide.

Entre los resultados principales del trabajo está la

Proposición 21. [El volumen de] cualquier segmento de un paraboloide de revolución es igual a la mitad del cono o segmento de un cono que tiene la misma base y el mismo eje.

La base es el área (fig. 5.3) de la figura plana, elipse o círculo, que se obtiene cortando el paraboloide por el plano que determina el segmento. La sección parabólica BAC y BC en la base son cortes mediante un plano que contiene al eje del paraboloide y es perpendicular al plano original. EF es la tangente a la parábola y por tanto, paralela a BC, y A es el punto de tangencia. AD, dibujado paralelo al eje del paraboloide, es el eje del segmento. Se puede demostrar que D es el punto medio de CB. Asimismo, si la base es una elipse, entonces CB es su eje mayor; si la base es un círculo, entonces CB es su

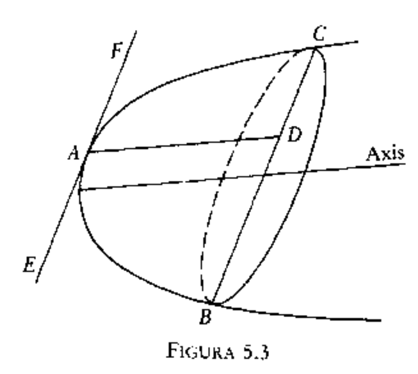


FIGURA 5.3

diámetro. El cono tiene la misma base que el segmento, vértice A y eje AD.

Proposición 24. Si a partir de un paraboloide de revolución se obtienen dos segmentos al cortar por dos planos *cualesquiera,* los volúmenes de los segmentos estarán en la misma razón que los cuadrados de los ejes respectivos.

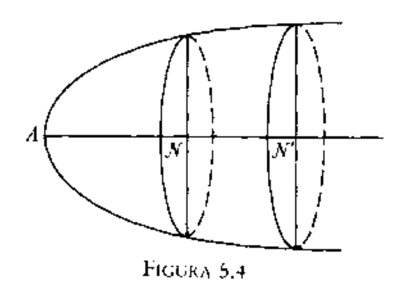


FIGURA 5.4

Para ilustrar el teorema, supongamos que los planos son perpendiculares al eje del paraboloide (fig. 5.4); entonces los dos volúmenes son uno al otro como AN^2 es a AN'^2. Hay teoremas semejantes para segmentos de hiperboloides y esferoides.

Uno de los trabajos más novedosos de Arquímedes es un corto tratado conocido como *El Método,* en el cual muestra cómo usó ideas procedentes de la mecánica para obtener teoremas matemáticos correctos. Este trabajo no fue descubierto hasta el año 1906 en una biblioteca de Constantinopla. El manuscrito está escrito durante el

siglo X en un pergamino que contiene otros trabajos de Arquímedes ya conocidos por otros caminos. Arquímedes ilustra su método de descubrimiento con el problema de encontrar el área de un segmento parabólico *CBA* (fig. 5.5). En el argumento, básicamente físico, usa teoremas sobre centros de gravedad ya establecidos por él.

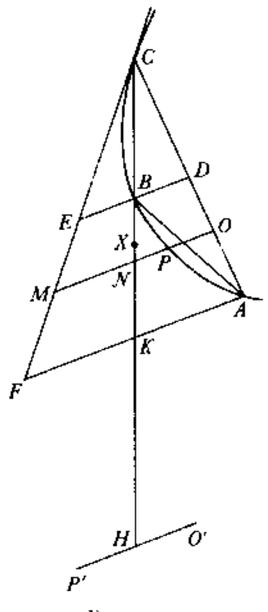


FIGURA 5.5

ABC (fig. 5.5) es un segmento arbitrario de una parábola limitado por la línea recta *AB* y el arco *ABC*. Sean *CE*, la tangente a la parábola en *C*; *D*, el punto medio de *CA*, y *DBE* el diámetro que contiene a *D* (línea paralela al eje de la parábola). Entonces Arquímedes afirma, tomando como referencia las *Cónicas* de Euclides, que

$$EB = BD, \qquad (1)$$

a pesar de que no se conoce la demostración de Euclides de este hecho. Se traza ahora el segmento *AF* paralelo a *ED* y sea *K* el punto de intersección de *CB* con *AF*. Determinamos el punto *H* en *CK* de

manera que $CK = KH$ y además, sea $MNPO$ un diámetro arbitrario de la parábola. Se tiene ahora, en virtud de (1) y el uso de triángulos semejantes, que $MN = NO$.

Arquímedes compara ahora el área del segmento y el área del triángulo CFA. Contempla la primera área como la suma de segmentos lineales tales como PO y el área del triángulo como la reunión de segmentos tales como MO. Prueba entonces que

$$HK \cdot OP = KN \cdot MO.$$

Desde el punto de vista físico esto significa que si consideramos KH y KN como los brazos de una palanca con el punto de apoyo en K, entonces OP considerado como un peso situado en H compensaría el peso MO situado en N. En consecuencia, colocando la suma de todos los segmentos lineales tales como PO en el punto H se compensará la suma de todos los segmentos lineales tales como MO, concentrado cada uno de ellos en su punto medio, que es el centro de gravedad de un segmento lineal. Pero la colección de segmentos MO, situado cada uno de ellos en su centro de gravedad, es «equivalente» al triángulo CAF situado en su centro de gravedad. En su libro *Sobre el Equilibrio de Planos,* Arquímedes prueba que este centro es el punto X situado en CK con $KX = (1/3)CK$. Por la ley de la palanca, $KX \cdot$ (área del triángulo CFA) $= HK \cdot$ (área del segmento parabólico), o bien

$$\frac{\triangle CFA}{\text{segmento } CBA} = \frac{HK}{KX} = \frac{3}{1} \tag{2}$$

Arquímedes deseaba relacionar el área del segmento con la del triángulo ABC. Concluye que (el área de) este triángulo es igual a la mitad de la del triángulo CKA puesto que ambos tienen la misma base CA y la altura de uno es la mitad de la altura del otro, como puede comprobarse con facilidad. Además, el triángulo CAK tiene un área igual a la mitad de la del triángulo CFA (ya que KA es la mitad del segmento FA). Luego el triángulo ABC tiene área igual a un cuarto del triángulo CFA y, por (2), se obtiene que el área del segmento ABC es al área del triángulo ABC como 4 es a 3.

En este método mecánico Arquímedes toma las áreas del segmento parabólico y del triángulo CFA como sumas de cantidades infinitas de segmentos lineales. Este método, en su opinión, lo es de descubri-

miento, pero no de demostración geométrica rigurosa. Prueba en este tratado que el uso de este procedimiento resulta eficaz a la hora de descubrir nuevos teoremas sobre esferas, cilindros, esferoides y paraboloides de revolución.

En su libro *Cuadratura de la Parábola,* Arquímedes da dos métodos para hallar el área de un segmento parabólico. El primero de ellos es semejante al argumento mecánico que acabamos de examinar y en el que de nuevo se compensan áreas mediante el principio de la palanca, pero su elección de las áreas es diferente. Su conclusión, naturalmente, coincide con (2) y se da en la proposición 16. Ahora, Arquímedes sabe el resultado que quiere probar y se dispone a hacerlo con rigor matemático a través de una sucesión de teoremas (proposiciones 18-24).

El primer paso es probar que el segmento parabólico puede «agotarse» mediante una serie de triángulos. Sea QPq (fig. 5.6a) el segmento parabólico y sea PV el diámetro que corta en dos partes iguales todas las cuerdas paralelas a la base Qq del segmento y de manera que V es el punto medio de Qq. Es intuitivamente claro, y se demuestra en la proposición 18, que la tangente en P es paralela a Qq. A continuación se toman QR y qS paralelos a PV y entonces el triángulo QPq es la mitad del paralelogramo $QRSq$, y así el triángulo QPq es mayor que la mitad del segmento parabólico.

Como corolario de este resultado, Arquímedes demuestra que el segmento parabólico se puede aproximar mediante un polígono tan cercano al mismo como se quiera, pues al construir un triángulo en el segmento limitado por PQ (fig. 5.6b), en el que P_1V_1 es el diámetro de ese segmento, se puede probar por métodos elementales de geometría (proposición 21) que (el área de) el triángulo $PP_1'q$, construido sobre Pq y que tiene las mismas propiedades que el triángulo PP_1Q, suman juntos 1/4 del triángulo PQq; además, en virtud del resultado del parágrafo anterior, los dos triángulos menores cubren más de la mitad de cada uno de los segmentos parabólicos en los que están situados. El proceso de construir triángulos sobre las nuevas cuerdas QP_1, P_1P, PP_1' y $P_1'q$ puede continuarse. Esta parte de la demostración es completamente análoga a la parte correspondiente en el teorema de Euclides sobre las áreas de dos círculos.

Así pues, tenemos condiciones suficientes para aplicar la proposición 1 del libro X de los *Elementos* de Euclides; es decir, podemos afirmar que el área de la figura poligonal obtenida al añadir triángulos al triángulo original PQq, es decir, el área

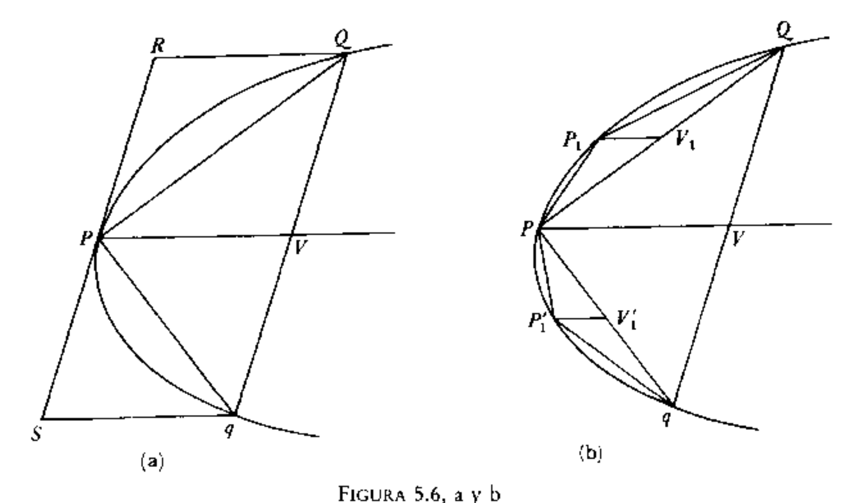


FIGURA 5.6, a y b

$$\triangle PQq + (1/4)\triangle PQq + (1/16)\triangle PQq + \ldots \qquad (3)$$

con una cantidad *finita* de términos se aproxima al segmento parabólico tanto como se quiera; esto es, la diferencia entre el área del segmento y la suma finita (3) puede hacerse menor que cualquier cantidad fijada previamente.

Arquímedes aplica ahora el método indirecto de demostración, que completa la prueba por el procedimiento de aproximaciones sucesivas. Demuestra en primer lugar que dados n términos de una progresión geométrica cuya razón es 1/4, se tiene

$$A_1 + A_2 + \ldots + A_n + (1/3)A_n = (4/3)A_1. \qquad (4)$$

Esto puede probarse con facilidad de varias maneras; podemos hacerlo con nuestra fórmula para la suma de n términos de una progresión geométrica. En la aplicación de (4), A_1 es el triángulo PQq.

Prueba entonces Arquímedes que el área A del segmento parabólico no puede ser ni mayor ni menor que $(4/3)A_1$. Su demostración consiste simplemente en que si el área A es mayor que $(4/3)A_1$ obtendría un conjunto (finito) de triángulos cuya suma S diferiría del

área del segmento en una cantidad menor que cualquier magnitud dada, por lo que la suma S sería mayor que $(4/3)A_1$. Así,

$$A > S > (4/3)A_1.$$

Pero por (4) si S contiene m términos, entonces

$$S + (1/3)A_m = (4/3)A_1$$

o bien

$$S < (4/3)A_1,$$

lo que es contradictorio.

Análogamente, supongamos que el área A del segmento parabólico es menor que $(4/3)A_1$. Entonces $(4/3)A_1 - A$ es un número positivo. Como los triángulos trazados por Arquímedes son cada vez más pequeños, podemos obtener una sucesión de triángulos inscritos tales que

$$(4/3)A_1 - A > A_m, \qquad (5)$$

donde A_m es el término m-ésimo de la sucesión y representa geométricamente la suma de 2^{m-1} triángulos. Pero como consecuencia de (4):

$$A_1 + A_2 + ... + A_m + \frac{1}{3}A_m = \frac{4}{3}A_1, \qquad (6)$$

entonces

$$\frac{4}{3}A_1 - (A_1 + A_2 + ... + A_m) = \frac{1}{3}A_m,$$

o bien

$$\frac{4}{3}A_1 - (A_1 + A_2 + ... + A_m) < A_m. \qquad (7)$$

Se sigue de (5) y (7) que

$$A_1 + A_2 + ... + A_m > A. \qquad (8)$$

Pero toda suma formada por triángulos inscritos es siempre menor que el área del segmento. Luego (8) es imposible.

Evidentemente, Arquímedes había sumado una progresión geométrica infinita, ya que cuando n tiene a infinito en (4), A_n tiende a cero, y la suma de la progresión infinita es $(4/3)A_1$.

Los trabajos de Arquímedes sobre los métodos mecánico y matemático de cálculo del área de un segmento parabólico ponen de manifiesto cómo distinguía con claridad entre los razonamientos físico y matemático. Su rigor es muy superior al que se puede encontrar en los trabajos de Newton y Leibniz.

En el trabajo *Sobre Espirales*, Arquímedes define la espiral como sigue: imaginemos que una línea (rayo) gira con velocidad angular constante alrededor de un extremo permaneciendo siempre en un mismo plano, y un punto que, comenzando por el extremo fijo, se mueve a lo largo de la línea con velocidad constante; entonces el punto describirá una espiral. En nuestras coordenadas polares la ecuación de la espiral es $\varrho = a\theta$. Tal como se ha dibujado la curva en la figura 5.7, θ se ha tomado en el sentido de las agujas del reloj. El resultado más profundo del trabajo es la

Proposición 24. El área limitada por la primera vuelta de la espiral y la línea inicial [el área sombreada en la figura] es igual a un tercio del primer círculo.

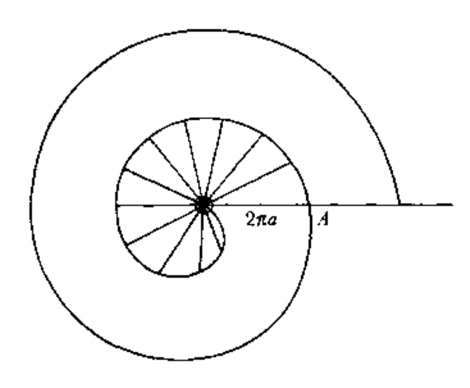


FIGURA 5.7

El primer círculo es el círculo de radio OA, que es igual a $2\pi a$, por lo que el área sombreada es $\pi(2\pi a)^2/3$.

La demostración se hace por el método de exhausción. En teoremas precedentes, en los que se preparan los instrumentos de demostración, el área de una región limitada por un arco de espiral, el arco $BPQRC$ de la figura 5.8, y por dos radios vectores OB y OC está contenida entre dos conjuntos de sectores circulares: así Bp', Pq', Qr', ... son arcos de círculos centrados en O y análogamente Pb, Qp, Rq son también arcos de círculos con centro en O. Los sectores circulares del conjunto inscrito son OBp', OPq', Oqr', ..., y los sectores circulares del conjunto circunscrito son OPb, OQp, ORq, ... Es decir, los sectores circulares sustituyen a los polígonos inscritos y circunscritos como figuras aproximadoras en el método de exhausción. (Nosotros utilizamos tales figuras en el cálculo cuando determinamos áreas en coordenadas polares.) La novedad de esta aplicación del método de las aproximaciones sucesivas es que Arquímedes elige sectores cada vez más pequeños de manera que la diferencia entre el área limitada por el arco de espiral y la suma de las áreas de la cantidad finita de sectores circulares «inscritos» (y la suma de las áreas de la cantidad finita de sectores circulares «circunscritos») se puede hacer menor que cualquier magnitud dada. Esta manera de aproximar el área no es la misma que «agotando» la misma añadiendo cada vez más figuras lineales. Sin embargo, en la última parte de la demostración Arquímedes utiliza el método indirecto de demostración igual que lo hace en el trabajo sobre la parábola y como lo hace Euclides en sus

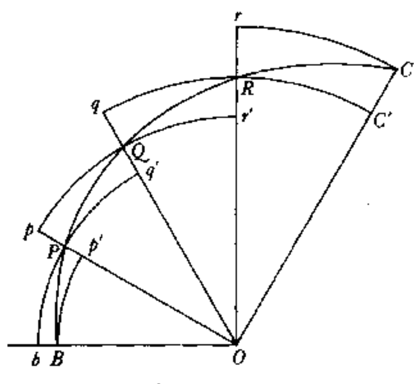


FIGURA 5.8

demostraciones por el método de las aproximaciones sucesivas. No hay ningún límite explícito en este proceso.

Arquímedes da también el resultado para el área limitada por el arco de espiral una vez que el radio vector ha dado dos vueltas completas alrededor de O; hay también otros resultados relacionados con áreas. Casualmente, matemáticos posteriores usaron la espiral para trisecar un ángulo y de hecho para dividir un ángulo en cualquier número de partes iguales.

Da la sensación de que, tras un estudio de sus trabajos geométricos, Arquímedes se dedicó exclusivamente en este campo a la obtención de resultados útiles sobre áreas y volúmenes. Estos trabajos, y sus trabajos matemáticos en general, no son espectaculares en cuanto a conclusiones, ni especialmente nuevos en cuanto a métodos o temas, pero aborda problemas muy difíciles y originales. Dice a menudo que las sugerencias de los problemas vienen de la lectura de los trabajos de sus predecesores; por ejemplo, los trabajos de Eudoxo sobre la pirámide, el cono y el cilindro (que aparecen en los *Elementos* de Euclides) sugirieron a Arquímedes su trabajo sobre la esfera y el cilindro, y la cuestión de la cuadratura del círculo sugirió la cuadratura del segmento parabólico. El trabajo de Arquímedes sobre hidrostática, no obstante, es completamente innovador; y sus trabajos sobre mecánica son nuevos en tanto que da demostraciones matemáticas (cap. 7, sec. 6). Su escritura es elegante, ordenada, acabada y a punto.

4. Areas y volúmenes en los trabajos de Herón

Herón, que vivió en algún momento entre los años 100 a. C. y 100 d. C., es de gran interés no sólo desde el punto de vista de la historia de las matemáticas, sino también para mostrar las características del período alejandrino. Proclo se refiere a Herón como *mecánico*, lo que podría significar un ingeniero mecánico de hoy y habla de él en conexión con Ctesibio, su maestro. Herón fue también un gran agrimensor.

Lo que más llama la atención de los trabajos de Herón es su mezcla de rigor matemático y lo aproximado de los métodos y fórmulas de los egipcios. Por otra parte, escribió un comentario sobre Euclides, usó los resultados precisos de Arquímedes (a los que se refiere con frecuencia), y en trabajos originales probó algunos teoremas nuevos de la geometría euclídea. Por otra parte, se dedicó a la

geometría aplicada y la mecánica y dio todo tipo de resultados aproximados sin justificación. Uso fórmulas egipcias con libertad y gran parte de su geometría fue también egipcia en cuanto a su carácter.

En sus *Métrica* y *Geométrica*, que han llegado hasta nosotros solamente a través de un libro que trata sobre su trabajo, Herón da teoremas y reglas para áreas planas, áreas de superficies y volúmenes de gran número de figuras. Los teoremas de estos libros no son nuevos. Para figuras con bordes curvilíneos utiliza los resultados de Arquímedes. Además, escribió *Geodesia* y *Estereometría* (cálculo de volúmenes de figuras), los cuales se refieren a las mismas cuestiones de los dos primeros libros. En todos estos trabajos está interesado principalmente en resultados numéricos.

En su *Dioptra* (teodolito), un tratado de geodesia, Herón muestra cómo calcular la distancia entre dos puntos de los que sólo uno es accesible y entre dos puntos visibles pero no accesibles. Muestra también cómo trazar una perpendicular desde un punto a una línea que no se puede alcanzar y cómo hallar el área de un campo sin entrar en él. La fórmula para el área de un triángulo, atribuida a él pese a ser debida a Arquímedes, es decir

$$\sqrt{s(s-a)(s-b)(s-c)},$$

donde a, b y c son los lados y s el semiperímetro, ilustra las ideas mencionadas con anterioridad. Esta fórmula aparece en la *Geodesia*, y la fórmula con una demostración está tanto en la *Dioptra* como en la *Métrica*. En la *Dioptra* muestra cómo excavar un túnel recto bajo una montaña trabajando simultáneamente desde ambos extremos.

Aunque algunas de sus fórmulas estan demostradas, Herón da varias sin demostración y otras son aproximadas. Así, da una fórmula inexacta para el área de un triángulo junto con la anterior correcta. Un motivo por el que Herón da varias fórmulas egipcias puede ser que las fórmulas exactas precisan raíces cuadradas o cúbicas y los agrimensores no ejecutaban tales operaciones. De hecho se distinguía entre geometría pura y geodesia o métrica. El cálculo de áreas y volúmenes pertenecía a la geodesia y no formaba parte de una educación general; estaba reservado a agrimensores, albañiles, carpinteros y otros técnicos. No hay ninguna duda de que Herón continuó y enriqueció la ciencia egipcia de la medida de campos; sus escritos sobre geodesia fueron utilizados durante varios siglos.

Herón aplicó varios de sus teoremas y reglas al diseño de teatros,

salas para banquetes y baños. Sus trabajos de aplicación incluyen *Mecánica, La Construcción de Catapultas, Mediciones, El Diseño de armas, Neumática* (la teoría y uso del aire comprimido), y *Sobre el Arte de Construcción de Autómatas.* Dio diseños para relojes de agua, instrumentos de medida, máquinas automáticas, máquinas elevadoras de pesos e ingenios de guerra.

5. Algunas curvas excepcionales

Pese a que los griegos clásicos introdujeron y estudiaron algunas curvas poco corrientes, como las cuadratrices, la máxima atención de esa geometría estuvo dedicada a figuras que podían dibujarse con regla y compás y relegó aquellas curvas al olvido. Los alejandrinos, sin embargo, se sintieron liberados de tal restricción; así Arquímedes no dudó en introducir la espiral. Varias curvas más fueron introducidas durante el período alejandrino.

Nicomedes (sobre el 200 a. C.) es conocido por su definición de la concoide. Comienza con un punto P y una línea AB (fig. 5.9); elige entonces una longitud a y coloca en todos los rayos que parten de P y cortan AB la longitud a partiendo del punto de intersección del rayo con AB, en la dirección que se aleja de P. Los puntos extremos así determinados son los puntos de la concoide. Así, los P_1, P_2 y P_3 de la figura son puntos de la concoide.

Si b es la distancia perpendicular de P a AB y si las longitudes a se miden a lo largo de los rayos que parten de P, y comenzando en AB pero en la dirección de P, obtenemos otras tres curvas según sea $a > b$, $a = b$ o $a < b$. Luego hay cuatro tipos de concoides, todas ellas debidas a Nicomedes. La ecuación polar moderna es $r = a + b \sec \theta$. Nicomedes usó la curva para trisecar un ángulo y duplicar el cubo [1].

Se atribuye a Nicomedes el invento de un mecanismo para construir las concoides. La naturaleza del mecanismo es de mucho menos interés que el hecho de que los matemáticos de la época estuvieran interesados en inventarlo. Las concoides de Nicomedes, junto con la recta y el círculo son las curvas constructibles más antiguas de las que poseemos una información satisfactoria.

---

[1] El método de trisección viene dado en T. L. Heath: *Los Trece Libros de los Elementos de Euclides.* Dover (reprint), 1956, vol. 1, p. 266.

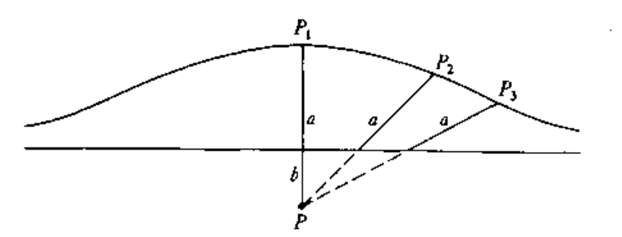


Figura 5.9

Diocles (final del siglo II a. C.), en su libro *Sobre los Espejos Ustorios* resuelve el problema de la duplicación del cubo introduciendo la curva llamada cisoide. La curva se define como sigue: *AB* y *CD* son diámetros perpendiculares de un círculo (fig. 5.10) y *EB* y *BZ* son arcos iguales. Se traza *ZH* perpendicular a *CD* y se traza entonces *ED*. La intersección de *ZH* y *ED* determina un punto *P* de la cisoide. Para Diocles la cisoide es el lugar geométrico de todos los puntos *P* determinados por todas las posiciones de *E* sobre el arco *BC* y *Z* sobre el arco *BD* con (arc *BE*) = (arc *BZ*). Se demuestra que

$$CH : HZ = HZ : HD = HD : HP.$$

Así *HZ* y *HD* son dos medias proporcionales entre *CH* y *HP*. Esto resuelve el problema de Delos. La ecuación de la cisoide en coordena-

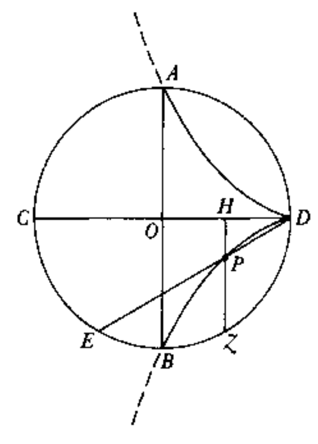


Figura 5.10

das rectangulares es $y^2(a + x) = (a - x)^3$, donde O es el origen; a el radio del círculo, y OD y OA los ejes de coordenadas. Esta ecuación incluye las dos ramas de la curva que se muestran en la figura, las cuales no fueron consideradas por Diocles.

6. El nacimiento de la trigonometría

Completamente nueva en la geometría cuantitativa griega alejandrina fue la trigonometría, una creación de Hiparco, Menelao y Ptolomeo. Este trabajo estuvo motivado por el deseo de construir una astronomía cuantitativa, y sería utilizada para predecir las trayectorias y posiciones de los cuerpos celestes y para ayudar a medir el tiempo, el cálculo del calendario, la navegación y la geografía.

La trigonometría de los griegos alejandrinos es lo que llamamos trigonometría esférica aunque, como veremos, incluye también las ideas básicas de la trigonometría plana. La trigonometría esférica presupone la geometría esférica, como por ejemplo las propiedades de los círculos máximos y los triángulos esféricos, muchas de las cuales ya eran conocidas; había sido investigada al mismo tiempo que la astronomía se convirtió en matemática, en los triángulos esféricos, muchas de las cuales ya eran conocidas; había sido investigada al mismo tiempo que la astronomía se convirtió en matemática, en los tiempos posteriores a los Pitagóricos. Los *Phaenomena,* de Euclides, basados asimismo en un antiguo trabajo, contienen algo de geometría esférica. Muchos de sus teoremas pretendían tratar sobre el movimiento aparente de las estrellas. Teodosio (sobre el 20 a. C.) recopiló los conocimientos aprovechables de entonces en su *Sphericae,* pero su trabajo no era numérico y por tanto no sería de utilidad para abordar el problema fundamental de la astronomía griega, es decir, medir el tiempo durante la noche mediante la observación de las estrellas.

El fundador de la trigonometría es Hiparco, que vivió en Rodas y Alejandría y murió alrededor del año 125 a. C. Conocemos muy poco acerca de él. La mayor parte de lo que conocemos proviene de Ptolomeo, que atribuye a Hiparco muchas ideas de trigonometría y astronomía. Le debemos a él varias observaciones astronómicas y descubrimientos, la teoría astronómica con mayor influencia en la antigüedad (cap. 7, sec. 4), y trabajos sobre geografía. De todos los trabajos de Hiparco solamente se ha conservado su *Comentario sobre los* Phaenomena *de Eudoxo y Aratus.* Gémino de Rodas escribió una

introducción a la astronomía, que poseemos, y que contiene una descripción del trabajo de Hiparco sobre el Sol.

El método de Hiparco de aproximarse a la trigonometría, como lo describió y utilizó Ptolomeo, es el siguiente. La circunferencia de un círculo se divide en 360°, tal como hizo por primera vez Hypsicles de Alejandría (sobre el 150 a. C.) en su libro *Sobre la Salida de los Astros* y por los babilonios de los últimos siglos antes de Jesucristo, y un diámetro se divide en 120 partes. Cada parte de la circunferencia y del diámetro se divide a su vez en 60 partes y cada una de ellas en otras 60, conforme al sistema babilónico de fracciones sexagesimales. Entonces, para un arco dado AB de un determinado número de grados, Hiparco —en un libro, perdido actualmente, sobre cuerdas en un círculo— da el número de unidades en la cuerda correspondiente AB. El método de cálculo de estas unidades será descrito en la exposición del trabajo de Ptolomeo, que presenta de manera combinada sus pensamientos y resultados.

El número de unidades de la cuerda correspondiente a un arco de un determinado número de grados equivale a la función seno moderna. Si 2α es el ángulo central del arco AB (fig. 5.11), para nosotros sen $\alpha = AC/OA$, mientras que, en vez de sen α, Hiparco da el número de unidades en $2 \cdot AC$ cuando el radio OA contiene 60 unidades. Por ejemplo, si la cuerda de 2α es de 40 unidades, para nosotros sen α = 20/60, o, con más generalidad,

$$\text{sen } \alpha = \frac{1}{60} \cdot \frac{1}{2} \text{ cuerda } 2\alpha = \frac{1}{120} \text{ cuerda } 2\alpha \qquad (9)$$

La trigonometría griega alcanzó una alta cota con Menelao (sobre

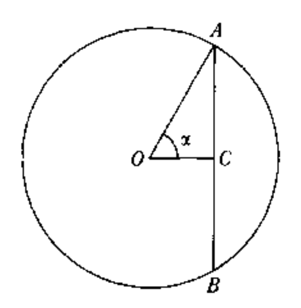


FIGURA 5.11

98 d. C.). Su *Sphaerica* es su obra capital, aunque parece ser que también escribió *Cuerdas en un Círculo* en seis libros y un tratado sobre la situación (o levantamiento) de arcos del Zodiaco. Los árabes le atribuyen algunas otras obras.

La *Sphaerica,* existente en versión árabe, está en tres libros. En el primero, sobre geometría esférica, se encuentra el concepto de triángulo esférico, es decir, la figura formada por tres arcos de círculos máximos sobre una esfera, cada uno de ellos menor que una semicircunferencia. El objetivo del libro es probar teoremas para triángulos esféricos, análogos a los probados por Euclides para los triángulos planos. Así, la suma de dos lados de un triángulo esférico es mayor que el tercer lado y la suma de los ángulos de un triángulo es mayor que dos ángulos rectos. Lados iguales abarcan ángulos iguales. Entonces Menelao demuestra el teorema, que no tiene análogo en los triángulos planos, según el cual si los ángulos de un triángulo esférico coinciden con los de otro, los dos triángulos son congruentes. Da también otros teoremas de congruencia y teoremas sobre triángulos isósceles.

El segundo libro de la *Sphaerica* de Menelao trata fundamentalmente de astronomía y sólo indirectamente se refiere a la geometría esférica. El tercer libro contiene algo de trigonometría esférica y bases para el desarrollo del primer teorema del libro, el cual supone que tenemos un triángulo esférico ABC (fig. 5.12) y algún círculo máximo que corta los lados del triángulo (trazado donde convenga). Para establecer el teorema usaremos nuestra moderna noción de seno, pero para Menelao el seno de un arco como AB (o el seno del ángulo central correspondiente en el centro de la esfera) se sustituye por la cuerda del arco doble AB. En términos de nuestro seno, el teorema de Menelao afirma que

$$\text{sen } P_1A \cdot \text{sen } P_2B \cdot \text{sen } P_3C = \text{sen } P_1\text{C} \cdot \text{sen } P_2A \cdot \text{sen } P_3B.$$

La demostración de este teorema se apoya sobre el teorema correspondiente para triángulos planos, llamado también teorema de Menelao. Para triángulos planos el teorema establece (fig. 5.13) que

$$P_1A \cdot P_2B \cdot P_3C = P_1C \cdot P_2A \cdot P_3B.$$

Menelao no demuestra el teorema plano. Se puede concluir que ya era

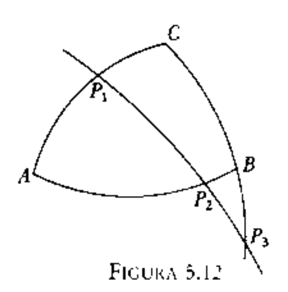


FIGURA 5.12

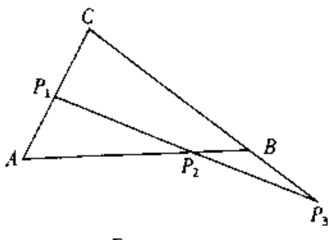


FIGURA 5.13

conocido o tal vez que Menelao lo había probado en un escrito anterior.

El segundo teorema del libro III, con la notación de que el arco a se opone al ángulo A en el triángulo ABC, dice que si ABC y $A'B'C'$ son dos triángulos esféricos y si $A = A'$ y $C = C'$ o C es suplementario de C', entonces

$$\frac{\text{sen } c}{\text{sen } a} = \frac{\text{sen } c'}{\text{sen } a'} .$$

El teorema 5 del libro III utiliza una propiedad de los arcos que era presumiblemente conocida en tiempos de Menelao, que es (fig. 5.14): si cuatro arcos de círculo máximo parten de un punto O y $ABCD$ y $A'B'C'D'$ son círculos máximos que cortan a los cuatro, se tiene:

$$\frac{\text{sen } AD}{\text{sen } DC} \cdot \frac{\text{sen } BC}{\text{sen } AB} = \frac{\text{sen } A'D'}{\text{sen } D'C'} \cdot \frac{\text{sen } B'C'}{\text{sen } A'B'} .$$

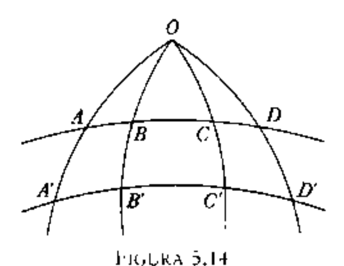


FIGURA 5.14

Encontraremos una expresión correspondiente a cada uno de los dos miembros reformulada bajo el concepto de razón anarmónica o razón doble en los trabajos de Pappus y en trabajos posteriores de geometría proyectiva. Se deben a Menelao muchos más teoremas sobre trigonometría esférica.

El desarrollo de la trigonometría griega y sus aplicaciones a la astronomía tuvieron su culminación en los trabajos del egipcio Claudio Ptolomeo (muerto el 168 a. C.), que era miembro de la familia real de matemáticos aunque no era de la casa real de Egipto. Ptolomeo vivió en Alejandría y trabajó en el Museo.

En su *Sintaxis Matemática* o *Colección Matemática* (el trabajo fue titulado por los árabes como *Megale Syntaxis, Megiste* y finalmente *Almagesto*), Ptolomeo continúa y completa los trabajos de Hiparco y Menelao en trigonometría y astronomía. La trigonometría y la astronomía están mezcladas en los trece libros del *Almagesto,* si bien el libro I trata con amplitud sobre trigonometría esférica y los restantes se dedican principalmente a la astronomía, de la que hablaremos en el capítulo 7.

El *Almagesto* de Ptolomeo es esencialmente matemático, salvo en los lugares en que utiliza la física aristotélica para refutar la hipótesis heliocéntrica, sugerida por Aristarco. Afirma que, debido a que solamente el conocimiento matemático, abordado interrogativamente, dará a sus practicantes un conocimiento fiable, había decidido cultivar tanto como le fuera posible esta disciplina teórica. Ptolomeo dice también que desea fundamentar su astronomía «sobre los caminos incontrovertibles de la aritmética y la geometría».

En el capítulo IX del libro I Ptolomeo comienza calculando las cuerdas de los arcos de un círculo, con lo que extendía los trabajos de Hiparco y Menelao. Como ya hemos observado, la circunferencia se divide en 360 partes o unidades (no usa la palabra «grado») y el diámetro en 120 unidades; propone entonces, dado un arco que contenga un determinado número de las 360 unidades, encontrar la longitud de la cuerda expresada en términos del número de unidades que contiene todo el diámetro, es decir, 120 unidades.

Comienza con el cálculo de las cuerdas de arcos de 36° y 72°. En la figura 5.15, ADC es un diámetro de un círculo con centro en D y BD es perpendicular a ADC. E es el punto medio de DC y F se elige de manera que $EF = BE$. Ptolomeo demuestra geométricamente que FD coincide con un lado del decágono regular inscrito y BF, con un lado del pentágono regular inscrito. Pero ED contiene 30 unidades y BD,

60 unidades. Como $EB^2 = ED^2 + BD^2$, $EB^2 = 4500$ y $EB = 67\ 4'55''$ (lo que representa $67 + 4/60 + 55/60^2$ unidades). Ahora, $EF = EB$ por lo que podemos conocer EF. Entonces $FD = EF - DE = 67\ 4'55'' - 30 = 37\ 4'55''$. Como FD es igual que el lado del decágono, es la cuerda de un arco de 36°. Luego conocemos la cuerda de este arco. Utilizando FD y el triángulo rectángulo FDB, podemos calcular BF: es igual a 70 32'3". Pero BF es el lado del pentágono, por lo que se tiene la cuerda del arco de 72°.

Naturalmente, para el lado de un hexágono regular, como coincide con el radio, se tiene evidentemente que la cuerda de longitud 60 pertenece al arco de longitud 60. Asimismo, como el lado del cuadrado inscrito se puede calcular de manera inmediata a partir del radio, se tiene la cuerda de 90°, que es 84 51'10". Además, puesto que el lado del triángulo equilátero inscrito puede calcularse también de manera inmediata a partir del radio, se obtiene que la cuerda de 120° es 103 55'23".

Con el uso del triángulo rectángulo ABC (fig. 5.16) sobre el diámetro AC se puede obtener inmediatamente la cuerda del arco suplementario AB si se conoce la cuerda del arco BC. Por tanto, como Ptolomeo conocía la cuerda de 36° podía calcular la de 144°, que resulta ser 114 7'37".

La relación que se ha establecido aquí es equivalente a $\text{sen}^2 A + \cos^2 A = 1$, donde A es un ángulo agudo arbitrario. Esto puede verse como sigue: Ptolomeo ha probado que si S es un arco menor de 180° entonces

$$(\text{cuerda } S)^2 + [\text{cuerda } (180 - S)]^2 = 120^2,$$

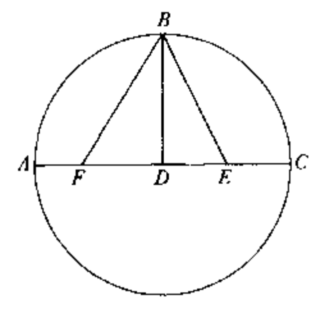


FIGURA 5.15

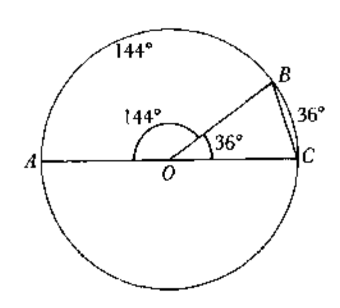


FIGURA 5.16

pero por la relación (9) anterior

$$(\text{cuerda } S)^2 = 120^2 \operatorname{sen}^2 \frac{S}{2}.$$

Luego se tiene

$$120^2 \operatorname{sen}^2 \frac{S}{2} + 120^2 \operatorname{sen}^2 \frac{180 - S}{2} = 120^2,$$

o bien

$$\operatorname{sen}^2 \frac{S}{2} + \operatorname{sen}^2 \left(90 - \frac{S}{2}\right) = 1,$$

es decir

$$\operatorname{sen}^2 \frac{S}{2} + \cos^2 \frac{S}{2} = 1.$$

Ahora Ptolomeo demuestra lo que él llama un lema, pero que se conoce hoy en día como el teorema de Ptolomeo: dado cualquier cuadrilátero inscrito en un círculo (fig. 5.17), demuestra que $AC \cdot BD = AB \cdot DC + AD \cdot BC$. La demostración es inmediata. Toma entonces el cuadrilátero especial $ABCD$ en el que AD es un diámetro (fig. 5.18). Supongamos que conocemos AB y AC. Ptolomeo muestra ahora cómo calcular BC. El segmento BD es la cuerda del arco

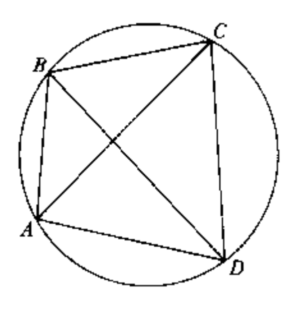


FIGURA 5.17

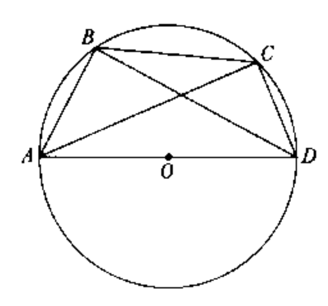


FIGURA 5.18

suplementario de AB, y CD es la cuerda del suplemento del arco AC. Si se aplica el lema, se ve que cinco de las seis longitudes involucradas en él son conocidas, por lo que la sexta, que en este caso es BC, se puede calcular. Pero (arco BC) = (arco AC) − (arco AB). Luego podemos calcular la cuerda de la diferencia de dos arcos cuando se conoce la cuerda de cada uno de ellos. Con la terminología moderna esto significa que si conocemos sen A y sen B podemos calcular sen $(A - B)$. Ptolomeo apunta que, puesto que conoce las cuerdas de 72° y 60°, puede calcular la de 12°.

Prueba a continuación cómo, dada una cuerda cualquiera en un círculo, se puede calcular la cuerda del arco mitad de la cuerda dada. En términos modernos esto representa calcular sen $A/2$ a partir de sen A. Este resultado es potente, como afirma Ptolomeo, ya que podemos comenzar con un arco cuya cuerda es conocida y calcular las cuerdas de sus sucesivas mitades. Prueba también que si se conocen las cuerdas de dos arcos AB y BC se puede calcular la cuerda del arco AC. Esto representa, en nuestro lenguaje actual, la fórmula de sen $(A + B)$. Como caso particular, se puede determinar, en términos modernos, sen $2A$ a partir de sen A.

Como Ptolomeo puede calcular la cuerda de 3/4° a partir de la cuerda de 12° mediante divisiones sucesivas en mitades, puede añadir este arco de 3/4° o restarlo de cualquier arco de cuerda conocida, y en virtud de los teoremas anteriores, puede calcular la cuerda de la suma o la diferencia de dos arcos. Por lo tanto, está en disposición de obtener las cuerdas de todos los arcos a intervalos de 3/4°. Sin embargo, desea obtener las cuerdas de arcos con saltos de 1/2°, lo que se dispone a hacer recurriendo a razonar con desigualdades. El resultado aproximado es que la cuerda de 1/2° es 0 31′25″.

Está ahora en disposición de construir una tabla de las cuerdas de arcos, para arcos que difieren entre sí 1/2°, desde 0° hasta 180°. Esta es la primera tabla trigonométrica.

Pasa entonces Ptolomeo (capítulo XI del libro I) a resolver problemas de astronomía, comenzando por encontrar arcos de círculos máximos sobre una esfera. Estos arcos son lados de triángulos esféricos, algunas de cuyas partes son conocidas bien por observación o mediante cálculos previos. Para determinar los arcos desconocidos, Ptolomeo prueba relaciones que son teoremas de trigonometría esférica, algunos de los cuales habían sido probados ya en el libro III de la *Sphaerica* de Menelao. El método básico de Ptolomeo consiste en usar el teorema de Menelao para triángulos esféricos. Así prueba, con

nuestra notación, que en el triángulo esférico con ángulo recto en C (fig. 5.19) y con arco a que denota el lado opuesto al ángulo A

$$\text{sen } a = \text{sen } c \text{ sen } A$$
$$\text{tg } a = \text{sen } b \text{ tg } A$$
$$\cos c = \cos a \cos b$$
$$\text{tg } b = \text{tg } c \cos A.$$

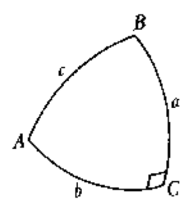


FIGURA 5.19

Por supuesto, para Ptolomeo las distintas funciones trigonométricas son cuerdas de arcos. Para tratar triángulos oblicuángulos los descompone en triángulos esféricos rectángulos. No hay ninguna presentación sistemática de la trigonometría esférica; demuestra únicamente aquellos teoremas que necesita para resolver problemas astronómicos concretos.

El *Almagesto* pone la trigonometría en su forma definitiva, que perdurará alrededor de mil años. Generalmente hablamos de esta trigonometría como esférica, pero la distinción entre trigonometría plana y esférica es muy difusa si se observa lo hecho por Ptolomeo. Ciertamente, Ptolomeo trabaja con triángulos esféricos pero, por haber calculado las cuerdas de arcos, ha puesto realmente las bases de la trigonometría plana. Pues, conociendo sen A y, por tanto, cos A para cualquier A comprendido entre 0° y 90°, se pueden resolver triángulos planos.

Obscrvemos que la trigonometría fue creada para ser usada en astronomía, y como la trigonometría esférica era de mayor utilidad para este propósito, fue la primera en ser desarrollada. El uso de la trigonometría plana en mediciones indirectas y en agrimensura es ajeno a la matemática griega. Esto puede parecernos extraño, pero es históricamente incuestionable, ya que la astronomía era el mayor objetivo de los matemáticos griegos. Los agrimensores hacen su

aparición en el período alejandrino; pero un matemático como Herón, que estuvo interesado en la agrimensura y habría sido capaz de desarrollar la trigonometría plana, se contentó con aplicar la geometría euclídea. Los agrimensores incultos no estaban en situación de crear la trigonometría necesaria.

7. La actividad geométrica tardía en Alejandría

La actividad matemática en general y geométrica en particular declinó en Alejandría aproximadamente a partir del comienzo de la era cristiana. Analizaremos las posibles razones del declive en el capítulo 8. Lo que sabemos acerca de los trabajos de geometría de la primitiva era cristiana viene de los principales comentaristas Pappus, Teón de Alejandría (fin del siglo IV d. C.) y Proclo.

En conjunto, muy pocos teoremas originales se descubrieron en este período. Los geómetras dan la impresión de haberse ocupado principalmente del estudio y comprensión de los trabajos de los grandes matemáticos que les precedieron. Completaron demostraciones que los autores originales habían omitido, bien porque las habían considerado suficientemente sencillas para dejarlas a los lectores, bien porque fueron dadas en tratados que se habían perdido. Estas demostraciones recibieron el nombre de lemas, en un antiguo uso de la palabra.

Tanto Teón como Pappus informan acerca de Zenodoro, que vivió en algún momento entre el 200 a. C. y el 100 d. C. Al parecer, Zenodoro escribió un libro sobre figuras isoperimétricas, es decir, figuras con el mismo perímetro y en él probó los teoremas siguientes:

1. Entre los polígonos de n lados con el mismo perímetro, el polígono regular es el que tiene mayor área.
2. Entre los polígonos regulares con igual perímetro, el que tiene más lados tiene mayor área.
3. El círculo tiene mayor área que un polígono regular del mismo perímetro.
4. De todos los sólidos con la misma superficie, la esfera tiene el mayor volumen.

El contenido de estos teoremas, que hoy en día llamaríamos problemas de máximos y mínimos, era novedoso en la matemática griega.

Al final del período alejandrino, las aportaciones de Pappus a la geometría aparecen como una especie de contrapunto. Los ocho libros de su *Colección Matemática* contienen algún material original. El nuevo trabajo de Pappus no fue de primer orden, pero algo del mismo merece ser tenido en cuenta.

El libro V da las demostraciones, resultados y extensiones de los trabajos de Zenodoro relativos a las áreas limitadas por curvas con el mismo perímetro. Pappus añade el teorema por el cual de todos los segmentos de un círculo que tienen el mismo perímetro, el semicírculo tiene mayor área. Prueba también que la esfera tiene mayor volumen que cualquier cono, cilindro o poliedro regular con la misma área de su superficie.

La proposición 129 del libro VII es un caso particular del teorema en el que la razón doble (fig. 5.20)

$$\frac{AB}{AD} \Big/ \frac{BC}{CD}$$

es la misma para toda sección transversal de cuatro rectas que parten de O. Pappus exige que las dos líneas transversales pasen por A.

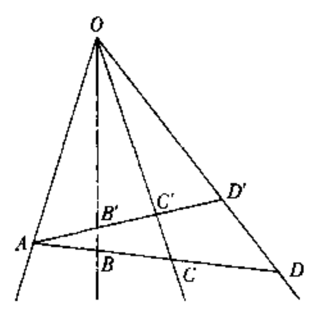


FIGURA 5.20

La proposición 130 afirma, en nuestro lenguaje, que si cinco de los puntos en los que los seis lados de un cuadrilátero completo (los cuatro lados y las dos diagonales) cortan una línea recta son fijos, el sexto también lo es. Así, si $ABCD$ (fig. 5.21) es un cuadrilátero tal que los seis puntos en los que sus seis lados cortan a una línea recta

arbitraria EK son E, F, G, H, J y K, si cinco de ellos son fijos, también lo es el sexto. Pappus observa que estos seis puntos verifican la condición

$$\frac{EK}{EH} \bigg/ \frac{JK}{JH} = \frac{EK}{EF} \bigg/ \frac{GK}{GF}.$$

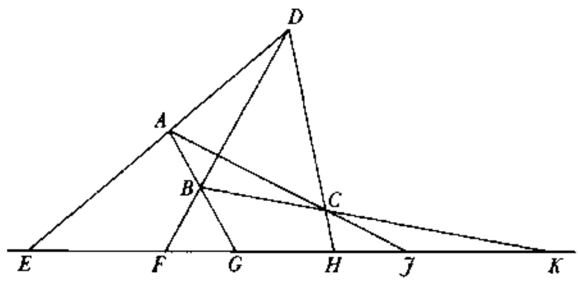


FIGURA 5.21

Esta condición establece que la razón doble determinada por E, K, J y H coincide con la razón doble determinada por E, K, G y F. La condición es equivalente a la que podemos encontrar, introducida por Desargues, que llama a seis puntos como los indicados, «puntos de una involución».

La proposición 131 del libro VII equivale a la afirmación de que la diagonal de cualquier cuadrilátero queda cortada armónicamente por la otra diagonal y por la línea que une los puntos de intersección de los pares de lados opuestos. Así, $ABCD$ es un cuadrilátero (fig. 5.22); CA es una diagonal; CA queda cortada por la otra diagonal BD y por FH, que une la intersección de AD y BC con la intersección de AB y CD. Entonces, los puntos C, E, A y G de la figura forman un conjunto armónico; es decir, E divide internamente a AC con la misma razón que G divide externamente a AC.

La proposición 139 del libro VII enuncia lo que se llama todavía teorema de Pappus. Si A, B y C son tres puntos de una recta (fig. 5.23) y A', B' y C' son tres puntos de otra, entonces AB', y $A'B$, BC' y $B'C$, y AC' y $A'C$ se cortan en tres puntos alineados.

Uno de los últimos lemas, la proposición 238, establece una propiedad fundamental de las secciones cónicas: el lugar geométrico

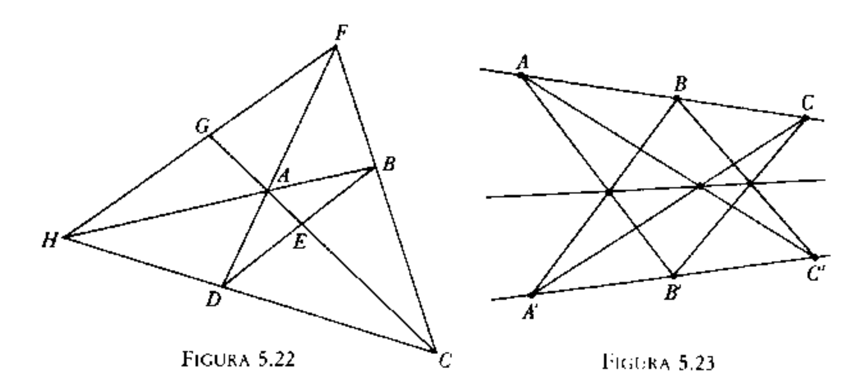


FIGURA 5.22 FIGURA 5.23

de todos los puntos cuyas distancias desde un punto fijo (foco) y desde una línea fija (directriz) están en razón constante es una sección cónica. Esta propiedad fundamental de las cónicas no aparece en el libro de Apolonio *Secciones Cónicas*, pero, como ya hemos observado en el capítulo precedente, era probablemente conocida por Euclides.

En la introducción del libro VII, Pappus se apoya en la afirmación de Apolonio de que su método capacita para hallar el lugar geométrico de los puntos tales que el producto de sus distancias a dos líneas es igual al producto de sus distancias a otras dos líneas por una constante. Pappus sabe —pero, sin embargo, no demuestra— que el lugar es una cónica. Apunta también que el problema se puede generalizar a cinco, seis o más rectas. Hablaremos de nuevo de esta cuestión en conexión con los trabajos de Descartes.

El libro VIII es de especial importancia puesto que está dedicado esencialmente a la mecánica, la cual, conforme a los puntos de vista alejandrinos, se contempla como una parte de la matemática. En efecto, Pappus prologa el libro planteando esta cuestión. Cita a Arquímedes, Herón y otras figuras menos conocidas como las figuras de la mecánica matemática. El centro de gravedad de un cuerpo se define como el punto interior del mismo (no ha de ser necesariamente interior) tal que si el cuerpo se suspende desde el mismo, permanece en su posición inicial. Explica entonces procedimientos para la determinación del punto. Trata también sobre el movimiento de un cuerpo a lo largo de un plano inclinado y aborda la cuestión de comparar la fuerza requerida para deslizar un cuerpo por un plano horizontal con la que se necesita para hacerlo en un plano inclinado.

El libro VII contiene también un famoso teorema llamado a veces teorema de Pappus y a veces teorema de Guldin, debido a que Paul Guldin (1577-1643) lo redescubrió de forma independiente. El teorema afirma que el volumen generado por la rotación completa de una curva cerrada plana totalmente situada a un lado del eje de rotación es igual al área limitada por la curva multiplicada por la circunferencia del círculo que pasa por el centro de gravedad. El resultado es muy general y Pappus era consciente de ello. No da una demostración del teorema y es muy posible que tanto el teorema como su demostración se conocieran con anterioridad a su tiempo.

En lo que se refiere a la geometría, el período alejandrino finaliza con los trabajos de varios comentaristas. Teón de Alejandría escribió un comentario sobre el *Almagesto* de Ptolomeo y nuevas ediciones de los *Elementos* y la *Optica* de Euclides. Su hija Hypatia (fallecida el 415), estudiante de matemáticas, escribió comentarios sobre Diofanto y Apolonio.

Proclo Diadoco, que hemos citado a menudo, escribió un comentario sobre el libro I de los *Elementos* de Euclides. Este comentario es importante porque Proclo había tenido acceso a trabajos ahora perdidos, incluyendo la *Historia de la Geometría* de Eudemo y el libro de Gémino que probablemente se titulaba la *Doctrina* o la *Teoría de la Matemática.*

Proclo recibió su educación en Alejandría y posteriormente se desplazó a Atenas, donde se convirtió en la cabeza de la Academia de Platón. Fue un avanzado neo-platónico y escribió varios libros sobre los trabajos de Platón y en general sobre filosofía; la poética ocupaba su interés tanto como la matemática. Igual que Platón, creía que la matemática era sierva de la filosofía. Es propedéutico, porque quiere limpiar los ojos del alma, eliminando los obstáculos que colocan los sentidos en el camino de los conocimientos universales.

Hubo otro lado no matemático de Proclo, que aceptó varios mitos y misterios religiosos y fue un devoto adorador de las divinidades griegas y orientales. Rechazó la teoría ptolemaica porque un caldeo «en quien no está permitido no creer» pensaba de manera distinta. Hay que decir que Proclo tuvo la suerte de que los oráculos caldeos no contradecían ni negaban a Euclides.

Entre otros comentaristas, citaremos unos pocos. Simplicio, un comentarista de Aristóteles, estudió en Alejandría y en la Academia de Platón, y se desplazó a Persia cuando Justiniano cerró la Academia el año 529. Reprodujo material de la *Historia* de Eudemo, incluyendo

un largo resumen del intento de Antifón sobre la cuadratura del círculo y sobre la cuadratura de lúnulas de Hipócrates. Isidoro de Mileto (siglo VI), que al parecer tuvo una escuela en Constantinopla (que se había convertido en la capital del Imperio Romano Oriental y el centro de alguna actividad matemática) escribió comentarios y puede haber escrito una parte del decimoquinto libro de los *Elementos* de Euclides. Eutocio (siglo VI d. C.), probablemente discípulo de Isidoro, escribió un comentario sobre los trabajos de Arquímedes.

Bibliografía

Aaboe, Asger: *Episodes from the Early History of Mathematics,* Random House, 1964, caps. 3-4.

Ball, W. W. R.: *A Short Account of the History of Mathematics,* Dover (reimpresión), 1960, caps. 4-6.

Cajori, Florian: *A History of Mathematics,* Macmillan, 1919, pp. 29-52.

Dijsterhuis, E. J.: *Archimedes* (Traducción Inglesa), Ejnar Munksgaard, 1956.

Heath, Thomas L.: *A History of Greek Mathematics,* Oxford University Press, 1921, vol. 2, caps. 13, 15, 17-19, 21.

—: *The Works of Archimedes,* Dover (reimpresión), 1953.

Pappus de Alejandría: *La Collection Mathématique,* ed. Paul Ver Eecke, 2 vols., Albert Blanchard, 1933.

Parsons, Edward Alexander: *The Alexandrian Library,* The Elsevier Press, 1952.

Sarton, George: *A History of Science*, Harvard University Press, 1959, vol. 2, caps. 1-3, 5, 18.

Scott, J. F.: *A History of Mathematics,* Taylor and Francis, 1958, caps. 3-4.

Smith, David Eugene: *History of mathematics,* Dover (reimpresión), 1958, vol. 1, cap. 4; vol. 2, caps. 8 y 10.

Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, caps. 7-8.

Capítulo 6

EL PERIODO ALEJANDRINO: EL RESURGIR DE LA ARITMETICA Y EL ALGEBRA

Dondequiera que haya un número está la belleza.

PROCLO

1. Los símbolos y operaciones de la aritmética griega

Volvamos por un momento a analizar las características de la aritmética en el período clásico. Los griegos clásicos llamaban *logística* al arte del cálculo; reservaban la palabra *aritmética* a la teoría de números. Los matemáticos clásicos desdeñaron la *logística* debido a que se refería a los cálculos prácticos que se necesitaban en la industria y el comercio. Sin embargo, nosotros consideraremos tanto la *logística* como la *aritmética* para ver lo que los griegos alejandrinos tenían a su disposición.

El arte clásico griego de escribir y trabajar con números no continúa en el lugar en que lo dejaron los babilonios. Parece que en *logística* los griegos se adaptaron a sus verdaderos orígenes. Los numerales arcaicos griegos encontrados en Creta anteceden al período clásico en aproximadamente quinientos años. No hay hechos notables en este esquema, solamente los símbolos concretos de los números 1, 2, 3, 4, 10, 200, 1000 y así sucesivamente. Muy al principio de la era clásica los griegos introdujeron otros símbolos especiales para los números y usaron algún tipo de ábaco para los cálculos. Posteriormente, alrededor del 500 a. C. usaron el sistema ático, del cual la

noticia más antigua que se tiene es una inscripción del año 450 a. C. El sistema ático usa palos para los números del 1 al 4; Π, la primera letra de *penta*, para el cinco, y más tarde se usó Γ para designar el 5; Δ, de *deka*, era el 10; H, de *hekaton*, representaba el 100; X, de *chilioi*, el 1000; y M, de *myrioi*, representaba el 10.000. Los números intermedios se representaban mediante combinaciones de estos símbolos especiales. Así Γ| = 6; ΓΔ = 50; ΓH = 5000; ΔΓ||| = 18.

No obstante, nadie sabe cómo escribían los números los primeros matemáticos clásicos —por ejemplo, los Pitagóricos—. Podían haber usado piedras para calcular, pues la palabra «cálculo» significa «piedra». El significado original en griego de «ábaco» era «arena», lo que sugiere que antes de la introducción de los ábacos, y probablemente más tarde dibujaban los números como marcas en la arena. En los trescientos años que van de Tales a Euclides los matemáticos no prestaron atención al cálculo, y este arte no progresó en absoluto. Es significativo que los libros no nos hablen de la práctica de la aritmética.

Por alguna razón desconocida, los griegos clásicos cambiaron su forma de escribir los números por el sistema jónico o el alejandrino, que usan letras del alfabeto. Este sistema alfabético era el más corriente entre los matemáticos greco-alejandrinos y se encuentra en particular en el *Almagesto* de Ptolomeo. Se usó también en la antigua Siria y en Israel.

Los detalles del sistema griego son como sigue:

α	β	γ	δ	ε	ς	ζ	η	θ
1	2	3	4	5	6	7	8	9
ι	κ	λ	μ	ν	ξ	ο	π	ο
10	20	30	40	50	60	70	80	90
ρ	σ	τ	υ	φ	χ	ψ	ω	T
100	200	300	400	500	600	700	800	900

Los números intermedios se escribían combinando los símbolos anteriores. Así $\iota\alpha = 11$, $\iota\beta = 12$, $\kappa\alpha = 21$ y $\varrho\nu\gamma = 153$. Los símbolos para el 6, 90 y 900 y el símbolo M del sistema ático no estaban en el alfabeto griego corriente de entonces; los tres primeros, llamados ahora stigma (o digamma), koppa y sampi, pertenecían a un antiguo alfabeto que los griegos habían tomado de los fenicios (quienes, no obstante, no utilizaban letras para designar números). El hecho de que

se usaran estas letras antiguas hace pensar que este sistema de escritura de los números data de aproximadamente el 800 a. C. y probablemente procede de Mileto, en Asia Menor.

Para números mayores de 1000, el alfabeto se repetía, pero se colocaba una coma antes de la letra para evitar confusiones. Asimismo, se trazaban líneas horizontales sobre los números para distinguirlos de las palabras. Por ejemplo:

$$,\overline{\alpha\tau\epsilon} = 1305.$$

Varios autores griegos usaban pequeñas variaciones del esquema anterior y de los esquemas dados más abajo.

Los papiros griegos de la primera parte del período alejandrino (los tres primeros siglos a. C.) contienen símbolos para el cero tales como $\overline{0} \cdot 0$, $\bar{0}$, $\overset{\frown}{0}$ y $\overline{0}$. El cero del período greco-alejandrino se usaba, igual que el cero del período seléucida-babilonio, para indicar unidades ausentes. Conforme a los manuscritos bizantinos, que son todo lo que tenemos de los trabajos de Ptolomeo, se usa el símbolo 0 para el cero tanto dentro como al final de un número.

El *Arenario* de Arquímedes presentaba un esquema para escribir números muy grandes. Intentaba demostrar que era capaz de escribir un número tan grande como los granos de arena del universo. Toma el mayor número expresado hasta entonces en numerales griegos, que es 10^8, una miríada de miríadas, y lo usa para comenzar una nueva serie de números que va hasta $10^8 \times 10^8$ ó 10^{16}. Utiliza entonces 10^{16} para iniciar una nueva serie de números de 10^{16} hasta 10^{24}, y así sucesivamente. A continuación hace una estimación del número de granos de arena del universo y demuestra que es menor que el mayor número que se puede expresar. Lo importante de este trabajo de Arquímedes no es un esquema práctico para escribir efectivamente cualquier número grande, sino la idea de que se pueden construir números grandes de manera indefinida. Apolonio tenía un esquema semejante.

Las operaciones aritméticas con los números enteros escritos y descritos antes eran semejantes a las nuestras. Así, para sumar, los griegos escribían los números uno debajo de otro para formar una columna de unidades, una de decenas, y así sucesivamente, sumando los números de cada columna y pasando de una columna a la siguiente. Estos métodos representaban un gran paso adelante si se comparan con los métodos egipcios. El último, sin embargo, fue utilizado también por los greco-alejandrinos.

Para las fracciones estaba el símbolo especial L'' para 1/2. Así (a veces con un acento) $\alpha L'' = 1\frac{1}{2}$, $\beta L'' = 2\frac{1}{2}$ y $\gamma L'' = 3\frac{1}{2}$. Las fracciones pequeñas se designaban escribiendo el numerador señalado con un acento y el denominador escrito una o dos veces cada una de ellas con dos acentos. Así, $\iota\gamma'\kappa\theta''\kappa\theta'' = 13/29$. Diofanto escribía a menudo el denominador antes que el numerador.

Se ha encontrado también el esquema egipcio para la escritura de fracciones cuyo numerador era mayor que 1 como una suma de fracciones unitarias. Así, Herón escribe $\frac{163}{224}$ como $\frac{1}{2}\frac{1}{7}\frac{1}{14}\frac{1}{112}\frac{1}{224}$, pero da también la expresión $\frac{1}{2}\frac{1}{8}\frac{1}{16}\frac{1}{32}\frac{1}{112}$ y otras expresiones de este tipo para la misma fracción. Usa también la forma griega anterior. De la misma forma, Ptolomeo escribe *algunas* fracciones como los egipcios, así $\frac{23}{25}$ como $\frac{1}{2}\frac{1}{3}\frac{1}{15}\frac{1}{50}$. El signo más fue siempre sobreentendido y, naturalmente, se usaban las letras del alfabeto griego donde nosotros utilizamos los numerales.

Las fracciones escritas tanto en el sistema griego como en el egipcio eran demasiado complicadas para los cálculos astronómicos, por lo que los astrónomos greco-alejandrinos adoptaron las fracciones sexagesimales babilonias. No se sabe con precisión cuándo comenzó esta práctica, pero se utiliza ya en el *Almagesto* de Ptolomeo. Así, cuando Ptolomeo escribe 31 25 quiere expresar $\frac{31}{60} + \frac{25}{60^2}$. Ptolomeo dice que usa las fracciones sexagesimales para evitar las dificultades de las fracciones ordinarias. Escribía los números *enteros* en base decimal pero no con la notación posicional. Sin embargo, los números enteros grandes aparecían tan raramente en los cálculos astronómicos que puede decirse que usaba la notación de posición sexagesimal. El uso del sistema sexagesimal de valor del lugar para las fracciones y de los numerales alfabéticos no posicionales para los números enteros parece peculiar e irracional. Sin embargo, nosotros escribimos todavía 130°15′17″.5.

Como se deduce de la discusión anterior, los alejandrinos usaron las fracciones como números en su verdadero sentido, mientras que los matemáticos del período clásico hablaban solamente de una razón entre números enteros y no como partes de un todo, y las razones se

utilizaban exclusivamente en las proporciones. Sin embargo, las fracciones genuinas, es decir, fracciones como entes con su verdadero significado, se usaron en el comercio incluso durante el período clásico. Durante el período alejandrino, Arquímedes, Herón, Diofanto y otros se sirvieron de las fracciones con entera libertad y efectuaron operaciones con ellas. Pese a todo, por lo que se puede saber, no trataron el concepto de fracción, al parecer debido a que es suficientemente claro desde el punto de vista intuitivo para que puedan ser aceptadas y utilizadas.

La raíz cuadrada como operación, aunque fue considerada en la Grecia clásica, pasó realmente desapercibida. Existen indicaciones en los escritos de Platón de que los pitagóricos aproximaban $\sqrt{2}$ sustituyendo 2 por 49/25 con lo que obtenían 7/5. Análogamente, Teodoro aproximó probablemente $\sqrt{3}$ tomando 49/16 en sustitución de 3, obteniendo 7/4. El número irracional como tal no tenía ningún lugar en la matemática de la Grecia clásica.

La siguiente información que tenemos acerca del manejo de raíces en Grecia procede de Arquímedes. En su *Medida del Círculo* aborda principalmente el cálculo de una buena aproximación de π, es decir, la razón de la circunferencia respecto al diámetro del círculo; en el transcurso del trabajo opera con números enteros grandes y con fracciones, y obtiene asimismo una excelente aproximación de $\sqrt{3}$, que es:

$$\frac{1351}{780} > \sqrt{3} > \frac{265}{153},$$

pero no da explicaciones acerca de cómo obtuvo este resultado. Entre las diversas conjeturas que aparecen en la literatura histórica respecto de su razonamiento la siguiente es muy plausible. Dado un número A, si lo escribimos como $a^2 \pm b$ donde a^2 es el cuadrado racional más próximo a A, mayor o menor, y b es el resto, entonces

$$a \pm \frac{b}{2a} > \sqrt{a^2 \pm b} > a \pm \frac{b}{2a \pm 1}.$$

El resultado de Arquímedes es obtenido mediante varias aplicaciones de este procedimiento. Para obtener la aproximación de π, Arquímedes demuestra en primer lugar que el área del círculo es igual

al área de un triángulo rectángulo cuya base tiene la misma longitud que la circunferencia del círculo y cuya altura coincide con el radio. Ahora, tiene que calcular la circunferencia. Esta la aproxima cada vez más usando polígonos regulares inscritos y circunscritos y calcula los perímetros de los mismos. Su resultado para π es:

$$3\frac{10}{71} < \pi < 3\frac{1}{7}.$$

Apolonio también escribió un libro sobre la cuadratura del círculo cuyo título es *Okytokion* (Distribución Rápida), en el que afirma haber mejorado la determinación de Arquímedes de π con métodos aritméticos más efectivos. Este es el único libro en el que Apolonio se aparta de los matemáticos griegos clásicos.

Herón aproximaba las raíces cuadradas con frecuencia mediante:

$$\sqrt{A} = \sqrt{a^2 \pm b} \sim a \pm \frac{b}{2a},$$

donde a y b tienen el mismo significado que antes. Obtiene esta aproximación tomando en primer lugar $\alpha = (c + A/c)/2$, donde c es algún valor determinado para $\sqrt{A}$; si escribimos A como $a^2 + b$ y tomamos $c = a$ entonces $\alpha = a + b/2a$. Herón mejora también el valor de α tomando $\alpha_1 = (\alpha + A/\alpha)/2$. Cuanto más próximo a $\sqrt{A}$ sea α, mejor será la aproximación α_1. La expresión fundamental de Herón para α la usaron también los babilonios.

Al final de la era alejandrina el algoritmo de la raíz cuadrada utiliza, igual que nosotros, el principio de que $(a + b)^2 = a^2 + 2ab + b^2$. Las aproximaciones sucesivas se obtienen por tanteo, teniendo cuidado siempre de que el cuadrado de la aproximación calculada sea menor que el número cuya raíz cuadrada se desea calcular. Al explicar el uso por parte de Ptolomeo de este procedimiento, Teón apunta que se utiliza una figura geométrica para ayudarse en la elección; esta figura es la que utiliza Euclides en la proposición 4 del libro II de los *Elementos* y es el camino geométrico de expresar $(a + b)^2$. Así, Ptolomeo da

$$\frac{103}{60} (+) \frac{55}{60^2} (+) \frac{23}{60^3}$$

para $\sqrt{3}$, lo que da el valor 1,7320509, que tiene seis cifras decimales exactas.

2. Crecimiento independiente de la aritmética y el álgebra

Hemos estado revisando los métodos para hacer aritmética empleados por los griegos en los dos períodos, pero más especialmente en el período alejandrino cuando la geometría y la trigonometría se convirtieron en materias de carácter cuantitativo. Pero el principal asunto a que se dedica este capítulo es el nacimiento de la aritmética y el álgebra como materias *independientes* de la geometría. Los trabajos aritméticos de Arquímedes, Apolonio y Ptolomeo fueron un paso en esta dirección, aunque usaron la aritmética para calcular cantidades geométricas. Se puede llegar a la conclusión de que los números estaban destinados a ello, debido a que representaban magnitudes geométricas y la lógica de las operaciones estaba garantizada por el álgebra geométrica. Pero no hay duda de que Herón, Nicómaco (sobre el 100), que fue probablemente un árabe de Gerasa en Judea, y Diofanto (sobre el 250), un griego de Alejandría, trataron los problemas aritméticos y algebraicos por sí mismos y no en dependencia de la geometría, ya sea como motivación, ya como auxiliar de la lógica.

Más significativo que el trabajo de Herón de calcular raíces cuadradas y cúbicas es el hecho de que formuló y resolvió problemas algebraicos mediante procedimientos aritméticos puros. No usaba símbolos especiales; la narración es verbal. Por ejemplo, trata el siguiente problema: dado un cuadrado tal que la suma de su área y su perímetro es 896 pies, determinar su lado. El problema, con nuestra notación, consiste en calcular x de manera que verifique $x^2 + 4x = 896$. Herón completa el cuadrado sumando 4 a cada miembro y tomando la raíz cuadrada. No demuestra nada, sino que simplemente describe las operaciones a realizar. Hay varios problemas de este tipo en sus trabajos. Evidentemente, éste es el viejo estilo egipcio y babilonio de presentación, y no hay ninguna duda que Herón recogió mucho material de los antiguos textos egipcios y babilonios. Allí, recordémoslo, el álgebra era independiente de la geometría y, como en el caso de Herón, una prolongación de la aritmética.

En su *Geometrica*, Herón habla de sumar un área, una circunferencia y un diámetro. Cuando usa estas palabras quiere decir, por supuesto, que lo que quiere es sumar sus valores numéricos. Del

mismo modo, cuando dice que multiplica un cuadrado por un cuadrado, quiere expresar que lo que está calculando es el producto de los valores numéricos respectivos. Herón tradujo también gran parte del álgebra geométrica de los griegos a procesos aritméticos y algebraicos.

Este trabajo de Herón (así como el uso que hace de las fórmulas egipcias de aproximación de áreas y volúmenes) se considera a veces como el principio del declive de la geometría griega. Es más correcto contemplarlo como una mejora helénica de las matemáticas babilonias y egipcias. Cuando Herón suma áreas y segmentos lineales, no está aplicando incorrectamente la geometría clásica griega sino que simplemente está continuando la práctica de los babilonios para quienes área y longitud eran exclusivamente palabras para ciertas incógnitas aritméticas.

Más notable desde el punto de vista del resurgimiento de una aritmética independiente es el trabajo de Nicómaco, que escribió la *Introductio Arithmetica* en dos tomos. Fue el primer libro de importancia en el que la aritmética (en el sentido de la teoría de números) estaba tratada con independencia absoluta de la geometría. Desde el punto de vista histórico, su importancia para la aritmética es comparable a la de los *Elementos* de Euclides para la geometría. Este libro no sólo fue estudiado, tomado como referencia y copiado por docenas de autores posteriores, sino que se reconoce como inspirador de varios libros por otros autores del mismo período, con lo que refleja el interés de la época. Los números representaban cantidades de objetos y dejaron ya de ser considerados como longitudes de líneas, como en Euclides. Nicómaco utiliza siempre palabras, mientras que Euclides emplea una letra, como A, o dos letras tales como BC —refiriéndose en el segundo caso a un segmento lineal— al hablar de números. Luego el lenguaje de Nicómaco es más torpe. Considera sólo números enteros y razones de números enteros.

Nicómaco era un pitagórico, y pese a que la tradición pitagórica no había muerto, la reanimó. De las cuatro materias destacadas por Platón —aritmética, geometría, música y astronomía— Nicómaco afirma que la aritmética es la madre de las demás. Esto es lo que mantiene:

no solamente porque decimos que existía antes que las demás en la mente del Dios creador como algún plan universal y ejemplar, confiando en ella como un diseño y arquetipo, el creador del universo puso en orden sus creaciones

materiales y las hizo de acuerdo con sus propios fines; sino también porque es por su naturaleza anterior en su nacimiento...

La aritmética, continúa, es esencial para las demás ciencias ya que éstas no existirían sin ella. Sin embargo, si las demás ciencias fueran abolidas, la aritmética seguiría existiendo.

La esencia de la *Introductio* está en los trabajos aritméticos de los primeros pitagóricos. Nicómaco considera números pares e impares, cuadrados, rectangulares y poligonales. Estudia también números primos y compuestos y números paralelepipédicos (de la forma $n^2(n + 1)$ y define muchos más tipos. Da la tabla de multiplicar para números comprendidos entre el 1 y el 9 precisamente como la aprendemos nosotros.

Nicómaco repite varios enunciados pitagóricos, como que la suma de dos números triangulares consecutivos es un cuadrado perfecto y recíprocamente. Va más allá que los pitagóricos al ver, aunque sin probarlas, relaciones de tipo general. Así, afirma, el $(n - 1)$-ésimo número triángular sumado con el número k-gonal n-ésimo da el número $(k + 1)$-gonal n-ésimo. Por ejemplo, el $(n - 1)$-ésimo número triangular sumado al número cuadrado n-ésimo da el n-ésimo número pentagonal. Con nuestra notación

$$\frac{n(n-1)}{2} + n^2 = \frac{n}{2}(3n - 1).$$

Asimismo, el n-ésimo número triángular, el n-ésimo número cuadrado, el n-ésimo número pentagonal, y así sucesivamente forman una progresión aritmética cuya diferencia es el $(n - 1)$-ésimo número triangular.

Descubrió la siguiente proposición: si escribimos los números impares:

$$1, 3, 5, 7, 9, 11, 13, 15, 17, \ldots$$

entonces el primero es el cubo de 1; la suma de los dos siguientes, el cubo de 2; la suma de los tres siguientes, el cubo de 3, y así sucesivamente. Hay otras proposiciones sobre progresiones.

Nicómaco da cuatro números perfectos, 6, 28, 496 y 8128 y repite la fórmula de Euclides para los mismos. Clasifica todo tipo de

razones, incluidas $m + 1 : m$, $2m + n : m + n$, y $mn + 1 : n$ y les da nombre. Estas fueron muy importantes en música.

Estudia también la proporción, la cual, dice, es muy necesaria para «las ciencias naturales, la música, la trigonometría esférica y la planimetría, y en especial para el estudio de la matemática antigua». Da diversos tipos de proporciones, entre ellas la proporción musical

$$a : \frac{a+b}{2} = \frac{2ab}{a+b} : b.$$

La *Introductio* da también la criba de Eratóstenes (de la que hablaremos en el capítulo 7); es un método para la obtención de números primos de manera rápida: se escriben todos los números impares a partir de 3 hasta donde se desee, entonces, se tachan todos los múltiplos de 3, es decir todos los números terceros mayores que 3. A continuación, se tachan todos los múltiplos de 5, o todos los números quintos mayores que 5, contando los que se puedan haber tachado ya. Luego, todos los números séptimos mayores que 7, y así sucesivamente. Se debe incluir ahora el 2 junto con los que no se han tachado. Estos son los números primos.

Nicómaco utiliza siempre números concretos para discutir las distintas categorías y proporciones. Los ejemplos ilustran y explican sus afirmaciones, pero no hay ningún apoyo tras los ejemplos para las afirmaciones generales. No utiliza el método deductivo de demostración.

La *Introductio* tuvo valor porque es una presentación sistemática, ordenada, clara y amplia de la aritmética de los enteros y las razones de enteros, liberada de la geometría. No era original en cuanto a las ideas pero fue una recopilación de gran utilidad. Incorporaba propiedades especulativas, estéticas, místicas y morales de los números, pero ninguna aplicación práctica. La *Introductio* fue el texto habitual de aritmética durante mil años. En Alejandría, a partir de la época de Nicómaco, la aritmética se convirtió en el tema de estudio favorito, por encima de la geometría.

En este tiempo, también el álgebra toma la delantera. Aparecen libros de problemas resueltos mediante técnicas algebraicas. Algunos de estos problemas eran exactamente los que aparecían en los textos babilonios del 2000 a. C. o en el papiro Rhind. Estos trabajos griegos sobre álgebra fueron escritos en forma literaria y no se usa ningún

simbolismo, ni tampoco se da ninguna demostración de los métodos empleados. A partir de la época de Nicómaco, los problemas que conducían a ecuaciones tenían la forma común de un rompecabezas. Entre cincuenta y sesenta de los mismos se conservan en el Códice Palatino de Epigramas Griegos (siglo X). Treinta de ellos como mínimo se atribuyen a Metrodoro (sobre el 500 d. C.), pero seguramente son más antiguos. Uno es el problema del ganado de Arquímedes, según el cual hay que calcular el número de bueyes y vacas de colores distintos conforme a la información dada. Otro se debe a Euclides e involucra a una mula y un asno que transportan grano. Otro se refiere al tiempo que deben emplear unas cañerías para llenar una cisterna. Había problemas de edades, tal como aparecen en nuestros textos de álgebra.

El punto culminante del álgebra greco-alejandrina se alcanza con Diofanto. No sabemos casi nada acerca de sus orígenes y de su vida; probablemente fue griego. Un problema algebraico hallado en una colección griega da los siguientes hechos acerca de su vida: su infancia duró 1/6 de su vida; su adolescencia, hasta 1/12 más; se casó tras 1/7 más, y su hijo nació 5 años después. El hijo vivió la mitad de la edad de su padre y éste murió 4 años después que el hijo. El problema es determinar cuánto vivió Diofanto. La respuesta se calcula con facilidad y resulta ser 84. Su trabajo destaca por encima del de sus contemporáneos; por desgracia, apareció demasiado tarde para tener una gran influencia en su tiempo, ya que una corriente de destrucción estaba engullendo toda la civilización.

Diofanto escribió varios libros que se han perdido en su totalidad. Se conoce parte de un tratado *Sobre los Números Poligonales* en el que establece y demuestra teoremas pertenecientes a los libros VII, VIII y IX de los *Elementos* con el método deductivo; sin embargo, los teoremas no son muy importantes. Su gran trabajo es la *Arithmetica,* la cual, dice Diofanto, comprende trece libros. Nosotros tenemos seis, procedentes de un manuscrito del siglo XIII que es una copia griega de otro más antiguo y de versiones posteriores.

La *Arithmetica,* como el papiro Rhind, es una colección de problemas independientes. La dedicatoria señala que fue escrito como una serie de ejercicios para ayudar a uno de sus estudiantes a aprender la materia. Uno de los hitos más importantes de Diofanto es la introducción del simbolismo en el álgebra. Debido a que no estamos en posesión del manuscrito escrito por él, sino de otro muy posterior, no conocemos los símbolos con exactitud. Se cree que el símbolo que

usaba para la indeterminada era ς, que jugaba el papel de nuestra x. Esta ς puede haber sido la misma letra que la griega σ escrita al final de una palabra, como en ἀριτμός (*arithmos*) y pudo haber sido elegida a causa de que no representaba ningún número en el sistema griego de servirse de letras para designar números. Diofanto llamó a la incógnita «el número del problema». Nuestra x^2 la escribe Diofanto como Δ^{Y}, la Δ por ser la primera letra de δύναμις (*dynamis*, «potencia»). x^3 es K^{Y}; la K de κύβος (*cubos*). x^4 es $\Delta^{Y}\Delta$; x^5 es ΔK^{Y}; x^6 es $K^{Y}K$. En este sistema K^{Y} no es exactamente el cubo de ς como x^3 lo es de x. Para Diofanto, $\varsigma^{x} = 1/x$. Usa también nombres para estas potencias, por ejemplo número para x, cuadrado para x^2, cubo para x^3, cuadrado-cadrado (dynamodynamis) para x^4, cuadrado-cubo para x^5 y cubo-cubo para x^6 [2].

La aparición de este simbolismo es evidentemente notable, pero el uso de potencias superiores a tres es todavía más extraordinario. Los griegos clásicos no podían ni querían considerar productos de más de tres factores ya que tal producto no tenía ningún significado geométrico. En una base puramente aritmética, no obstante, tales productos tienen un significado; y ésta es precisamente la idea adoptada por Diofanto.

Diofanto indica la adición poniendo los términos uno a continuación de otro. Así

$$\Delta^{Y}\overline{\gamma}\mathring{M}\iota\overline{\beta} \quad \text{significa} \quad x^2 \cdot 3 + 12.$$

La $\mathring{M}$ es un símbolo para la unidad e indica que a continuación va un número puro que no contiene la indeterminada. De nuevo

$$\Delta^{Y}\overline{\alpha}\varsigma\overline{\beta}\mathring{M}\overline{\gamma} \quad \text{significa} \quad x^2 + x \cdot 2 + 3.$$

Para la sustracción emplea el símbolo ⩚. Así, para indicar $x^6 - 5x^4 + x^2 - 3x - 2$ escribe:

$$K^{Y}\kappa\overline{\alpha} \pitchfork \Delta^{Y}\Delta\overline{\varepsilon}\varsigma\overline{\gamma}\mathring{M}\overline{\beta},$$

poniendo todos los términos negativos detrás de los positivos. No existen símbolos para la adición, la multiplicación y la división como

[2] Algunos autores modernos usan δ, K y $\overline{\nu}$ en vez de Δ, K e Y.

operaciones. El símbolo ι^{σ} se emplea (al menos en las versiones existentes de la *Arithmetica*) para designar la igualdad. Los coeficientes de las expresiones algebraicas son números concretos; no hay ningún símbolo para coeficientes generales. A causa del uso de símbolos, el álgebra de Diofanto recibe el nombre de sincopada, mientras que a la de los egipcios, los babilonios, Herón y Nicómaco se le llama retórica.

Diofanto redacta sus soluciones en un texto continuo, de la misma manera que nosotros escribimos prosa. Su ejecución de las operaciones es completamente aritmética; es decir, no hay ninguna llamada a la geometría para ilustrar o justificar sus afirmaciones. Así $(x - 1)(x - 2)$ se interpreta algebraicamente, igual que lo hacemos nosotros. Aplica también identidades algebraicas tales como

$$\left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2 = pq$$

a expresiones tales como $x + 2$ en sustitución de p y $x + 3$ por q. Esto es, da pasos en los que utiliza las identidades pero estas identidades propiamente dichas no aparecen.

El primer libro de la *Arithmetica* consiste fundamentalmente en problemas que conducen a la determinación de ecuaciones de primer grado con una o más incógnitas. Los cinco libros restantes tratan principalmente de ecuaciones indeterminadas de segundo grado. Pero esta separación no es estricta. En el caso de ecuaciones determinadas (es decir, ecuaciones con solución única) con más de una incógnita, emplea la información dada para eliminar todas las incógnitas menos una y, al final, termina con ecuaciones cuadráticas de la forma $ax^2 = b$. Por ejemplo, el problema 27 del libro I dice: encontrar dos números tales que su suma sea 20 y su producto 96. Diofanto procede así: sea 20 la suma; 96, el producto, y $2x$, la diferencia entre los números buscados. Luego, los números son $10 + x$, $10 - x$. Por tanto, $100 - x^2 = 96$. Entonces, $x = 2$ y los números buscados son 12 y 8.

La característica más sorprendente del álgebra de Diofanto es su solución de las ecuaciones indeterminadas. Tales ecuaciones habían sido consideradas con anterioridad, como por ejemplo en el trabajo pitagórico para las soluciones de $x^2 + y^2 = z^2$, así como en el problema arquimediano del ganado, que conduce a siete ecuaciones con ocho incógnitas (más dos condiciones suplementarias), y en otros

escritos curiosos. Diofanto, no obstante, trata las ecuaciones indeterminadas extensamente y es el creador de esta rama del álgebra llamada en la actualidad, efectivamente, análisis diofántico.

Resuelve ecuaciones lineales con dos incógnitas, como

$$x + y - 5 = 0.$$

En estas ecuaciones da un valor a una indeterminada y resuelve la ecuación para un valor racional positivo de la otra. Reconoce que el valor asignado a la primera incógnita es estrictamente accidental. (En el análisis diofántico moderno solamente se calculan soluciones enteras.) Muy poco es lo hecho con este tipo de ecuación y el trabajo es apenas significativo puesto que las soluciones racionales positivas se calculan de golpe.

Resuelve entonces ecuaciones cuadráticas, cuya forma más general es (en nuestra notación)

$$y^2 = Ax^2 + Bx + C. \qquad (1)$$

Diofanto no escribe y^2 pero dice que la expresión cuadrática debe ser igual a un número cuadrado (cuadrado de un número racional). Considera (1) para valores especiales de A, B y C y estudia estos tipos en casos separados. Por ejemplo, cuando no aparece C toma $y = mx/n$, donde m y n son números enteros concretos, obtiene

$$Ax^2 + Bx = \frac{m^2}{n^2}x^2,$$

y entonces simplifica x y la resuelve. Cuando A y C no se anulan pero $A = a^2$, supone $y = ax - m$. Si $C = c^2$, pone $y = (mx - c)$. En todos los casos m es un número determinado.

Estudia también el caso de ecuaciones cuadráticas simultáneas, como

$$y^2 = Ax^2 + Bx + C \qquad (2)$$

$$z^2 = Dx^2 + Ex + F. \qquad (3)$$

Aquí también considera solamente casos particulares, es decir, cuando $A, B, \ldots, F$ son números determinados o verifican condiciones especia-

les, y este método consiste en asignar a y y a z expresiones en términos de x, y a continuación resolver en x.

De hecho, está resolviendo ecuaciones determinadas en una incógnita. Lo que ocurre, sin embargo, es que al elegir expresiones para y y z en (2) y (3) y para y en (1), está dando soluciones totalmente significativas y los valores asignados a y y z son arbitrarios.

Tiene también problemas en los que expresiones cúbicas y de mayor grado de x deben ser iguales a un número cuadrado, por ejemplo:

$$Ax^3 + Bx^2 + Cx + d^2 = y^2.$$

Aquí toma $y = mx + d$ y fija m de manera que se anule el coeficiente de x. Como el término d^2 se simplifica y puede dividir por x^2 toda la ecuación, obtiene una ecuación de primer grado en x. Hay también casos especiales en que una expresión cuadrática en x es igual a y^3.

Reduce todas estas ecuaciones cuadráticas en x a los tipos

$$ax^2 = bx,\ ax^2 = b,\ ax^2 + bx = c,\ ax^2 + c = bx,\ ax^2 = bx + c$$

y resuelve cada uno de ellos. Solamente resuelve una ecuación cúbica en x, de escasa importancia.

Las ecuaciones anteriores muestran los *tipos* de problemas que resuelve Diofanto. El lenguaje real de los problemas se ilustra con los ejemplos siguientes:

Libro I, problema 8. Dividir un número cuadrado dado en dos cuadrados.

Aquí toma el 16 como el número cuadrado dado y obtiene 256/25 y 144/25. Este es el problema que generalizó Fermat y que da lugar a la afirmación de que la ecuación $x^m + y^m = z^m$ no es resoluble para $m > 2$.

Libro II, problema 9. Dividir un número dado que es la suma de dos cuadrados en la suma de otros cuadrados distintos de los anteriores.

Toma 13 ó 4 + 9 como número dado y obtiene 324/25 y 1/25.

Libro III, problema 6. Encontrar tres números tales que su suma y la suma de dos cualesquiera de ellos sea un cuadrado perfecto.

Diofanto da los tres números 80, 320 y 41.

Libro IV, problema 1. Dividir un número dado en dos cubos tales que la suma de sus lados es un número dado.

Con el número 370 y la suma de los lados igual a 10, encuentra el 43 y el 27. Los lados son las raíces cúbicas de los cubos.

Libro IV, problema 29. Expresar un número dado como la suma de cuatro cuadrados más la suma de sus lados.

Dado el número 12, halla 121/100, 49/100, 361/100 y 169/100 como los cuatro cuadrados; sus lados son las raíces cuadradas de cada cuadrado.

En el libro VI Diofanto resuelve varios problemas que hacen referencia a los lados (racionales) de un triángulo rectángulo. El uso del lenguaje geométrico es accidental, incluso cuando aparece el término área. Lo que busca son números racionales *a*, *b* y *c* tales que $a^2 + b^2 = c^2$ y sujetos a alguna otra condición. Así, el primer problema es hallar un triángulo rectángulo (de lados racionales) tal que la hipotenusa menos cada uno de los catetos es un cubo. En este caso llega a obtener la solución entera 40,96 y 104. Sin embargo, en general obtiene soluciones racionales.

Diofanto da muestras de una gran habilidad para reducir ecuaciones de los diferentes tipos a formas que pueda manejar. No sabemos cómo llegó a sus métodos. Como prescinde totalmente de la geometría, no es probable que se inspirara en los procedimientos empleados por Euclides para resolver ecuaciones cuadráticas. Además, Euclides no estudia problemas indeterminados, que, como tales, aparecen con Diofanto. Debido a que carecemos de información acerca de la continuidad del pensamiento en los finales del período alejandrino, no podemos hallar indicios de los trabajos de Diofanto en sus predecesores. Dentro de lo que se puede asegurar, sus trabajos en álgebra pura son notablemente distintos de trabajos precedentes.

Acepta solamente raíces racionales positivas e ignora cualesquiera otras. Incluso cuando una ecuación cuadrática tiene dos raíces positivas da solamente una, la mayor. Cuando al resolver una ecuación observa claramente que va a dar dos raíces negativas o imaginarias, la rechaza y dice que no es resoluble. En el caso de raíces irracionales, recompone sus pasos y prueba cómo modificando la ecuación se puede obtener otra con raíces racionales. Aquí Diofanto se aparta de Herón y Arquímedes. Herón era un agrimensor y las cantidades geométricas que buscaba podían ser irracionales, por lo que las aceptaba, aunque naturalmente las aproximaba para obtener valores

útiles. Arquímedes buscaba también soluciones exactas, y cuando eran irracionales determinaba desigualdades para acotarlas. Diofanto es un algebrista puro, y como el álgebra de su tiempo no admitía los números irracionales, ni los negativos, ni los complejos, rechazaba las ecuaciones que tenían tales soluciones. Es, no obstante, digno de destacar que las fracciones, para Diofanto, eran números y no una razón entre dos números enteros.

No tenía ningún *método general.* Cada uno de los 189 problemas de la *Arithmetica* está resuelto por un procedimiento distinto. Hay más de 50 tipos diferentes de problemas, pero no se hace ningún intento de clasificarlos. Sus métodos son más cercanos a los de los babilonios que a los de sus antecesores griegos y hay indicios de influencias babilónicas, puesto que, en efecto, resuelve algunos problemas como lo hacían los babilonios. Sin embargo, no se ha probado la existencia de una conexión directa entre los trabajos de Diofanto y el álgebra babilonia. Sus adelantos en álgebra respecto de los babilonios consisten en la utilización del simbolismo y la resolución de ecuaciones indeterminadas. En ecuaciones determinadas no hizo mucho más que ellos, pero su *Arithmetica* se parecía a la *logística,* que Platón, entre otros, había proscrito de la matemática.

La variedad de métodos de Diofanto para cada uno de los problemas deslumbra más que deleita. Fue un virtuoso sagaz y clarividente pero en apariencia no profundizó lo suficiente en sus procedimientos para darles generalidad. (Es también cierto que el análisis diofántico es un conjunto de problemas independientes.) Al contrario de un investigador especulativo que busca ideas generales, Diofanto buscaba solamente soluciones correctas. Hay algunos resultados que podrían llamarse generales, como que ningún número primo de la forma $4n + 3$ puede ser la suma de dos cuadrados. Euler atribuyó a Diofanto métodos generales del mismo que no parecían tales debido a que no tenía coeficientes literales, y otros creyeron que su material pertenecía a una ciencia abstracta y básica. Pero este punto de vista no fue compartido por todos. Sin embargo, sus trabajos, contemplados como un todo, son un monumento algebraico.

Un elemento de la matemática que es de gran importancia hoy en día, y que faltaba en el álgebra griega, es el uso de letras para representar una clase de números como, por ejemplo, los coeficientes de las ecuaciones. Aristóteles utilizó letras del alfabeto griego para indicar un tiempo arbitrario o una distancia cualquiera y en la discusión del movimiento empleaba frases tales como «la mitad de *B*».

Euclides, asimismo, usaba letras para designar clases de números en los libros VII a IX de los *Elementos*, una práctica continuada por Pappus. Sin embargo, no hubo ningún reconocimiento de la enorme contribución que podían aportar las letras para aumentar la efectividad y la generalidad de la metodología algebraica.

Otra característica del álgebra alejandrina es la ausencia de una estructura deductiva explícita. Los diferentes tipos de números —números enteros, fracciones e irracionales— no estaban realmente definidos. No existía ninguna base axiomática sobre la que se pudiera levantar una estructura deductiva. Los trabajos de Herón, Nicómaco y Diofanto, y de Arquímedes en lo que concierne a la aritmética, siguen la metodología de los textos egipcios y babilonios que dicen cómo hacer las cosas. Las demostraciones deductivas y ordenadas de Euclides y Apolonio, y de la geometría de Arquímedes, han sido olvidadas. Los problemas son inductivos en espíritu ya que muestran métodos para problemas concretos que presumiblemente pueden ser aplicados a clases generales cuya amplitud no se específica. A la vista del hecho de que, como consecuencia de los trabajos de los matemáticos de la Grecia clásica, se admitían determinados resultados obtenidos deductivamente a partir de una base axiomática explícita, la aparición de una aritmética y un álgebra independientes sin ninguna estructura lógica de su edificio se convirtió en uno de los grandes problemas de la historia de la matemática. Este acercamiento a la aritmética y el álgebra es el indicio más claro de las influencias egipcias y babilonias en el mundo alejandrino. Pese a que los algebristas greco-alejandrinos no parecieron conscientes de estas deficiencias, veremos que iban a preocupar profundamente a los matemáticos europeos.

Bibliografía

Ball, W. W. R.: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, caps. 5 y 7.

Cajori, Florian: *A History of Mathematics*, Macmillan, 1919, pp. 52-62.

Heath, Thomas L.: *Diophantus of Alexandria*, Dover (reimpresión), 1964.

—: *The Works of Archimedes*, Dover (reimpresión), 1953, caps. 4 y 6 de la introducción, pp. 91-98, 319-326.

—: *A History of Greek Mathematics*, Oxford University Press, 1921, vol. 1, caps. 1-3; vol. 2, cap. 20.

—: *A Manual of Greek Mathematics,* Dover (reimpresión), 1963, caps. 2-3 y 17.

D'Ooge, Martin Luther: *Nichomachus of Gerasa,* University of Michigan Press, 1938.

Van der Waerden, B. L.: *Science Awakening*, P. Noordhoff, 1954, pp. 278-86.

Capítulo 7
LA RACIONALIZACION GRIEGA DE LA NATURALEZA

La matemática es la puerta y la llave de las ciencias.

ROGER BACON

1. La inspiración de la matemática griega

Desgraciadamente, salvo indicaciones ocasionales, los clásicos griegos, como los *Elementos* de Euclides, las *Secciones Cónicas* de Apolonio y los trabajos geométricos de Arquímedes, no dan ninguna explicación de por qué estos autores investigaron sus temas. Dan solamente la matemática formal y cuidadosamente deductiva. En este sentido, los textos griegos no presentan diferencias con los modernos manuales y tratados de matemáticas. Tales libros pretenden exclusivamente organizar y presentar los resultados matemáticos alcanzados y por ello omiten las motivaciones de la matemática, las pistas y sugerencias de los teoremas y las aplicaciones del conocimiento matemático.

Para comprender por qué los griegos crearon una matemática de gran vitalidad, se deben investigar sus objetivos. Lo que impulsó a los griegos a crear y apreciar la matemática fue el deseo urgente e irreprimible de comprender el mundo físico. La matemática era una parte importante de la investigación de la Naturaleza y la clave para la comprensión del Universo, pues las leyes matemáticas son la esencia de su diseño.

¿Qué pruebas tenemos de que éste fue el papel de la matemática? Resulta difícil demostrar que un teorema determinado o un conjunto de teoremas fueran creados para una finalidad concreta puesto que no tenemos información suficiente acerca de los matemáticos griegos. La afirmación directa de Ptolomeo de que creó la trigonometría para la astronomía es una excepción. Sin embargo, cuando se descubre que Eudoxo fue primordialmente un astrónomo y que Euclides no escribió solamente los *Elementos* sino también *Phaenomena* (un trabajo sobre la geometría de la esfera aplicado al movimiento de la esfera celeste), la *Optica y Catóptrica*, los *Elementos de Música*, y pequeños trabajos sobre mecánica, todos ellos matemáticos, no podemos por menos de concluir que las matemáticas fueron algo más que una disciplina aislada. Conociendo cómo funciona la mente humana, y conociendo con gran detalle cómo trabajaban hombres como Euler y Gauss, podemos estar seguros de que las investigaciones en astronomía, óptica y música sugirieron problemas matemáticos y es más que probable que la motivación para la matemática fuese su aplicación a estas otras áreas. Es asimismo relevante que la geometría de la esfera, conocida en Grecia como «sphaerica» fuera estudiada al mismo tiempo que la Astronomía se convertía en matemática, lo que ocurrió incluso antes de los tiempos de Eudoxo. La palabra «sphaerica» significaba «astronomía» para los pitagóricos.

Afortunadamente, las deducciones que podemos hacer a partir de los trabajos de los matemáticos, pese a ser bastante razonables, se establecen sin duda por la evidencia abrumadora de los escritos de los filósofos griegos, muchos de los cuales eran también eminentes matemáticos, y de los científicos griegos. Los límites de la matemática no eran propiamente matematicos. En el período clásico, la matemática comprendía la aritmética, la geometría, la astronomía y la música, mientras que en el período alejandrino, como ya hemos observado en el capítulo 5, las secciones de las ciencias matemáticas eran aritmética (Teoría de Números), geometría, mecánica, astronomía, óptica, geodesia, canónica (armonía musical) y logística (aritmética aplicada).

2. Los comienzos de una visión racional de la naturaleza

Las civilizaciones que precedieron a la griega o las que eran contemporáneas de ella contemplaban la naturaleza como algo caótico, misterioso, caprichoso y terrorífico. Los acontecimientos de la

naturaleza estaban manipulados por dioses. Sacerdotes y magos podían inducir a los dioses a ser bondadosos e incluso a realizar milagros, pero la vida de los hombres estaba completamente sometida a su antojo.

A partir de la época en que nuestro conocimiento de la civilización y cultura griegas comienza a ser razonablemente definido y concreto, es decir, alrededor del año 600 a. C., encontramos en los intelectuales una actitud completamente nueva frente a la naturaleza: racional, crítica y laica. La mitología fue rechazada, al creer que los dioses manipulaban al hombre y al mundo físico de acuerdo con sus caprichos. La nueva doctrina establece que la naturaleza es ordenada y funciona invariablemente conforme a un plan. Además, se tiene la absoluta convicción de que la mente humana es potente e incluso superior; el hombre no sólo puede aprender los caminos de la naturaleza, sino que incluso puede predecir los acontecimientos.

Es cierto que la aproximación racional fue asumida solamente por los intelectuales, un grupo pequeño tanto en el período clásico como en el alejandrino. Mientras estos hombres eran contrarios a la atribución de los acontecimientos a dioses y demonios y desafiaban los misterios y terrores de la naturaleza, el pueblo en general era profundamente religioso y creía que los dioses dominaban los sucesos. Aceptaban doctrinas místicas y supersticiones crédulamente igual que lo hacían egipcios y babilonios. De hecho, la mitología griega fue amplia y altamente desarrollada.

Los jonios comenzaron la tarea de determinar la naturaleza de realidad. No describiremos las teorías cualitativas de Tales, Anaxágoras y sus compañeros, cada uno de ellos centrado en una sustancia concreta inmutable a través de todos los cambios aparentes. La identidad subyacente de esta sustancia primera se conserva, pero todas las formas de la materia pueden explicarse en términos de ella. Esta filosofía natural de los jonios dio lugar a una serie de denodadas especulaciones, audaces conjeturas y brillantes intuiciones más bien que resultados de amplias y cuidadosas investigaciones científicas. Fueron quizá excesivamente ambiciosos queriendo ver la escena en su totalidad y dando el salto ingenuamente para sacar conclusiones. Pero ellos sustituyeron las viejas historias míticas por explicaciones materiales y objetivas de la estructura y el diseño del universo. Ofrecieron un acercamiento razonable en lugar de las narraciones imaginativas e ingenuas de los poetas y defendían sus afirmaciones con el uso de la razón. Al menos estos hombres se atrevieron a estudiar el universo

con sus mentes y rechazaron la confianza en dioses, espíritus, fantasmas, demonios, ángeles y demás agentes míticos.

3. El desarrollo de la creencia en una estructura matemática

El paso decisivo para eliminar el misterio, el misticismo y la arbitrariedad de los trabajos sobre la naturaleza, y reducir la apariencia de caos a un modelo comprensible y ordenado fue la aplicación de la matemática. El primer grupo importante que presentó una filosofía racional y matemática de la naturaleza fue el de los pitagóricos. Sacaron alguna inspiración a partir del aspecto místico de la religión griega; sus doctrinas religiosas estaban centradas en la purificación del alma y su redención desde la corrupción y encierro del cuerpo. Sus componentes vivían con sencillez y estaban dedicados por entero al estudio de la filosofía, la ciencia y la matemática. Los nuevos miembros debían prometer el secreto, al menos en las creencias religiosas, y les exigían formar parte del grupo durante toda la vida. La comunidad estaba abierta tanto a los hombres como a las mujeres.

El pensamiento religioso de los pitagóricos era indudablemente místico, pero su filosofía natural era decididamente racional. Estaban sorprendidos por el hecho de que fenómenos que eran de muy diferente forma desde el punto de vista cualitativo, presentaban propiedades matemáticas idénticas. Por lo tanto, las propiedades matemáticas debían ser la esencia de tales fenómenos. Más concretamente, los pitagóricos basaban esta esencia en el número y en las relaciones numéricas. El número era su primer principio para la explicación de la naturaleza. Todos los objetos estaban formados por puntos o «unidades de existencia» en combinaciones que correspondían a las distintas figuras geométricas. Como pensaban en los números a la vez como puntos y como partículas elementales de materia, el número era la materia y la forma del Universo y la causa de todo fenómeno. De aquí la doctrina pitagórica «todas las cosas son números». Dice Filolao, un famoso pitagórico del siglo V: «si no fuera por el número y su naturaleza, nada de lo que existe estaría claro para nadie, ni en sí mismo ni en su relación con otras cosas. Podéis observar la potencia del número influyendo no sólo en los negocios de demonios y dioses, sino también en todos los actos y en el pensamiento del hombre, en todos los oficios y en la música».

La reducción de la música, por ejemplo, a una simple relación

entre números fue posible para los pitagóricos cuando descubrieron dos hechos: primero, que el sonido producido por una cuerda pulsada depende de la longitud de la cuerda; y segundo, que los sonidos armónicos están dados por cuerdas igualmente tirantes cuyas longitudes respectivas estén en razón igual a las razones de números enteros. Por ejemplo, un sonido armónico se produce pulsando dos cuerdas igualmente tensas si la longitud de una de ellas es igual al doble de la de la otra. En nuestro lenguaje, el intervalo entre las dos notas es una octava. Otra combinación armónica está formada por dos cuerdas cuyas longitudes están en la relación de 3 a 2, en este caso la más corta da una nota llamada la quinta inferior de la dada por la primera cuerda. De hecho, las longitudes relativas en toda combinación armónica de cuerdas pulsadas se pueden expresar como una razón de números enteros. Los pitagóricos desarrollaron también una famosa escala musical griega. Pese a que no dedicaremos espacio a la música en el período griego, hagamos notar que varios matemáticos griegos, incluidos Euclides y Ptolomeo, escribieron sobre el tema, en especial acerca de las combinaciones armónicas de sonidos y la construcción de escalas.

Los pitagóricos redujeron los movimientos de los planetas a relaciones numéricas. Creían que los cuerpos, al moverse en el espacio, producían sonidos; tal vez esto les vino sugerido por el zumbido que se produce al blandir un objeto sujeto al extremo de una cuerda. Creían, además, que un cuerpo que se mueve con rapidez da una nota más alta que el que se mueve lentamente. Ahora, de acuerdo con su astronomía, un planeta se mueve más rápidamente cuanto mayor es su distancia a la Tierra. Luego los sonidos que producen los planetas, que no oímos porque estamos acostumbrados a ellos desde el nacimiento, variaban con su distancia a la Tierra y todos están armonizados. Pero como esta «música de las esferas», como toda armonía, no se reducía más que a relaciones numéricas, lo mismo ocurría con los movimientos de los planetas.

Los pitagóricos, y probablemente el propio Pitágoras, buscaban no solamente observar y describir los movimientos celestes, sino también encontrar alguna regularidad en ellos. La idea de un movimiento circular uniforme, aparentemente obvia en el caso de la Luna y el Sol, sugería que todos los movimientos planetarios fueran explicables en términos de movimientos circulares uniformes. Los últimos pitagóricos llevaron a cabo una ruptura más que sorprendente con la tradición; fueron los primeros en creer que la Tierra era esférica.

Además, como el 10 era su número ideal, decidieron que debía haber 10 cuerpos en movimeinto en el cielo. Primero, existía un fuego central alrededor del cual se movían los cuerpos celestes, *incluida la Tierra*. Conocían cinco planetas además de la Tierra. Estos seis cuerpos, el Sol, la Luna y la esfera a la que estaban sujetas las estrellas daban un total de 9 cuerpos móviles solamente. En consecuencia, afirmaban la existencia de un décimo cuerpo, llamado Anti-Tierra, que también giraba alrededor del fuego central. Nosotros no podemos ver este décimo cuerpo debido a que se mueve exactamente a la misma velocidad que la Tierra en el lado opuesto del fuego central y también porque la parte habitada de la Tierra está de espaldas al fuego central. Aquí tenemos la primera teoría que pone la Tierra en movimiento. Sin embargo, los pitagóricos no afirmaban la rotación de la esfera; al contrario, la esfera de estrellas fijas gira alrededor del centro del Universo.

La creencia en que los cuerpos celestes son eternos, divinos, perfectos e inmutables y que los objetos sublunares, como la Tierra y (de acuerdo con los griegos) los cometas, están sujetos a cambios, descomposición, decadencia y muerte, puede haber venido también de los pitagóricos. La doctrina del movimiento circular uniforme y la distinción entre cuerpos celestes y sublunares se convirtió en algo inherente al pensamiento griego.

Otros hechos de la naturaleza fueron «reducidos» también a números. Los números 1, 2, 3 y 4, la *tetractys*, fueron especialmente apreciados porque su suma es 10. De hecho el juramento de los pitagóricos parece haber sido: «Yo juro en nombre de la Tetractys que ha sido concedida a nuestra alma. La fuente y las raíces de todo lo que fluye en la naturaleza están contenidos en ella.» Los pitagóricos afirmaban que la naturaleza estaba formada de agrupaciones de cuatro elementos; por ejemplo, punto, línea, superficie y sólido, y los cuatro elementos, tierra, aire, fuego y agua. Los cuatro elementos fueron también centrales en la filosofía de la naturaleza de Platón. Como el 10 era ideal, el 10 representaba el Universo. La idealidad del 10 exigía que la totalidad del Universo se pudiera describir en términos de 10 categorías de opuestos: impar y par, limitado e ilimitado, bueno y malo, derecho e izquierdo, uno y varios, macho y hembra, recto y curvo, cuadrado y oblongo, luz y oscuridad, y reposo y movimiento.

Evidentemente, la filosofía pitagórica mezcló pensamientos serios con doctrinas que podríamos considerar fantasiosas, inútiles y sin base científica. Su obsesión por la importancia de los números dio

como resultado una filosofía natural que en realidad tenía poca correspondencia con la naturaleza. No obstante, orientaron la comprensión de la naturaleza no como los jonios, a través de una sustancia única, sino por medio de la estructura formal de relaciones numéricas. Además, tanto ellos como los jonios decían que el verdadero sentido de los datos debía ser un orden armonioso de la naturaleza.

Podemos ver ahora por qué el descubrimiento de cantidades inconmensurables fue algo desastroso para la filosofía pitagórica: una razón de longitudes inconmensurables no se podría expresar como una razón de números enteros. Además, creían que la recta estaba formada por una cantidad finita de puntos (que identificaban con partículas físicas); pero éste no sería el caso para una longitud como $\sqrt{2}$. Su filosofía, basada en la primacía de los números enteros, habría saltado hecha añicos si hubieran aceptado los irracionales como números.

Puesto que los Pitagóricos «reducían» la astronomía y la música a números, estas disciplinas quedaron enlazadas a la aritmética y la geometría, y las cuatro fueron consideradas como las disciplinas matemáticas. Se convirtieron y permanecieron en parte del *currículum* escolar incluso en la Edad Media, cuando recibieron el nombre colectivo de «el *quadrivium*». Como ya hemos observado, el interés de los pitagóricos por la aritmética (es decir, por la teoría de números) no fue debido a los valores puramente estéticos de dicha disciplina sino a una búsqueda de la explicación de fenómenos naturales en términos numéricos; y esta valoración puso énfasis en proporciones especiales y en las formas triangulares, cuadrangulares, pentagonales y de orden superior en que se pueden ordenar los números. Además, fue la filosofía natural de los pitagóricos, centrada alrededor del número, la que dio importancia a esta materia con autores como Nicómaco. De hecho, la ciencia moderna coincide con el énfasis de los pitagóricos sobre el número —aunque, como veremos, de una manera mucho más sofisticada— mientras que la teoría moderna de números, puramente estética, proviene de la aritmética pitagórica propiamente dicha.

Los filósofos situados cronológicamente entre los pitagóricos y Platón estudiaron igualmente la naturaleza de la realidad, pero no involucraron en ella la matemática de una manera directa. Los argumentos y puntos de vista de hombres como Parménides (siglo V a. C.), Zenón (siglo V a. C.), Empédocles (*c.* 484-*c.* 424 a. C.), Leucipo (*c.* 440 a. C.) y Demócrito (*c.* 460-*c.* 370 a. C.) fueron cualitativos,

igual que los de sus predecesores jonios. Hicieron grandes afirmaciones acerca de la realidad, que eran, en el mejor de los casos, escasamente sugeridas por la observación. Sin embargo, todos afirmaban que la naturaleza es inteligible y que la realidad puede entenderse a través del pensamiento. Cada uno de ellos era un eslabón de la cadena que conducía a la investigación matemática de la naturaleza. Leucipo y Demócrito fueron notables porque fueron los más explícitos al afirmar la teoría del atomismo. Su filosofía común era que el mundo está compuesto de una cantidad infinita de átomos simples y eternos. Estos difieren en la forma, tamaño, orden y posición, pero todo objeto es una combinación de tales átomos. Pese a que las magnitudes geométricas se pueden dividir indefinidamente, los átomos son las últimas partículas indivisibles (la palabra *átomo* significa indivisible en griego). Dureza, forma y tamaño son propiedades físicamente reales de los átomos. Las restantes propiedades, como el sabor, calor y color no están en los átomos sino en el perceptor; así el conocimiento sensorial es imprevisible porque varía con el perceptor. Igual que los Pitagóricos, los atomistas aseguraban que la realidad que subyace en la constantemente cambiante diversidad del mundo físico se podía expresar en términos matemáticos y, además, que los acontecimientos de este mundo estaban estrictamente determinados por leyes matemáticas.

Platón, el primero de los pitagóricos después de Pitágoras, fue el propagador con mayor influencia de la doctrina por la que la realidad e inteligibilidad del mundo físico se pueden abarcar solamente a través de las matemáticas. Para él no había ninguna duda de que el mundo estaba matemáticamente trazado, ya que «Dios geometriza eternamente». El mundo percibido por los sentidos es confuso y engañoso y en cualquier caso imperfecto y perecedero. El conocimiento físico no es importante, ya que los objetos materiales cambian y decaen; así, el estudio directo de la naturaleza y las investigaciones estrictamente físicas son inútiles. No obstante, el mundo físico es una copia imperfecta del mundo ideal, el único que deben estudiar matemáticos y filósofos. Las leyes matemáticas, eternas e inmutables, son la esencia de la realidad.

Platón fue más allá que Pitágoras al desear no solamente comprender la naturaleza a través de la matemática, sino sustituir la propia naturaleza por ella. Creía que alguna ojeada al mundo físico podría suministrar algunas verdades básicas, a partir de las cuales la razón podría seguir adelante sin ayuda. Desde esa óptica no existiría la

naturaleza, sino la matemática, que sustituiría las investigaciones físicas como lo hace en geometría.

La actitud de Platón hacia la astronomía ilustra su postura sobre la búsqueda del conocimiento. Esta ciencia no se refiere al movimiento de los cuerpos celestes visibles. La ordenación de las estrellas en el cielo y sus movimientos aparentes son por supuesto maravillosos y bellos para ser contemplados, pero la simple observación y explicación de los movimientos se apartan de la verdadera astronomía. Antes de llegar a esta última «debemos dejar solo al cielo», pues la verdadera astronomía trata de las leyes del movimiento de las estrellas verdaderas en un cielo matemático del que el cielo visible no es más que una expresión imperfecta. Platón anima a una dedicación a una astronomía teórica, cuyos problemas agradan a la mente y no a los ojos, y cuyos objetos se abarcan mentalmente y no con la vista. La variedad de figuras que presenta el cielo a los ojos se utiliza sólo como diagramas para ayudar a la investigación de las verdades más altas. Los usos de la astronomía en la navegación, cálculo del calendario y la medida del tiempo fueron ajenos a Platón.

Los puntos de vista de Platón acerca del papel de la matemática son una parte integral de su filosofía, que afirma que hay una realidad objetiva universalmente válida constituida de formas e ideas. Estas realidades eran independientes de los seres humanos y eran inmutables, eternas e intemporales. Estas ideas han llegado hasta nosotros a través del recuerdo o anamnesis; si bien están presentes en el alma, deben estimularse para reparar en ellas o hacerlas surgir desde sus abismos. Estas ideas son la única realidad. Incluidas en ellas, pero ocupando un rango inferior, están las ideas matemáticas, que se contemplan como un estado intermedio entre el mundo sensible y las ideas superiores como la bondad, la verdad, la justicia y la belleza. En esta filosofía general la matemática jugaba un doble papel; no sólo eran una parte de la realidad sino que, como ya hemos señalado en el capítulo 3, ayudaban a disciplinar a la mente para alcanzar las ideas eternas. Como dice Platón en el libro VII de *La República*, el estudio de la geometría da una visión más sencilla del concepto de divinidad: «La geometría conduce el alma hacia la verdad y crea el espíritu de la Filosofía...»

Aristóteles, al mismo tiempo que tomaba varias ideas de su maestro Platón, tenía una idea completamente distinta del estudio del mundo real y de la relación entre las matemáticas y la realidad. Criticó la visión del mundo de Platón y su reducción de la ciencia a las

matemáticas. Aristóteles fue un físico; creía en las cosas materiales como sustancia primera y origen de la realidad. La física y la ciencia en general debe estudiar el mundo físico para obtener verdades; el verdadero conocimiento se obtiene a partir de la experiencia sensorial por medio de la intuición y la abstracción. Entonces, la razón debe aplicarse a los conocimientos así obtenidos.

La materia por sí misma no es importante. Como tal es indeterminada, simplemente tiene la potencialidad de la forma; la materia se convierte en algo importante cuando está organizada en formas diversas. La forma y el cambio en la materia, que dan lugar a nuevas formas, son los hechos interesantes de la realidad y los que ciertamente conciernen a la ciencia.

Al contrario de lo que creían los antiguos griegos, Aristóteles pensaba que la materia no estaba formada de una sustancia primitiva. La materia que vemos y tocamos está compuesta de cuatro elementos básicos: tierra, agua, fuego y aire. A su vez, cada elemento tiene sus propias cualidades características. La tierra es fría y seca; el agua, fría y húmeda; el aire, caliente y húmedo, y el fuego, caliente y seco. Por tanto, las cualidades de un objeto determinado dependen de las proporciones de los elementos que entran en él; y con esto quedan determinadas la solidez, la dureza, el grosor y otras propiedades.

Los cuatro elementos tienen otras características. La tierra y el agua tienen gravedad; el aire y el fuego, ligereza. La gravedad motiva que un elemento tienda a situarse en el centro de la Tierra; la ligereza da lugar a que busque el cielo. Así, si se conocen las proporciones de los elementos que forman parte de un cuerpo dado, se puede conocer también su movimiento.

Aristóteles contempla los sólidos, fluidos y gases como tres tipos distintos de materia que se distinguen por la posesión de diferentes cualidades intrínsecas. La transición de sólido a líquido, por ejemplo, representa la pérdida de una cualidad y su sustitución por otra. Así, el cambio del mercurio en oro rígido conlleva tomar del mercurio la sustancia que tiene la fluidez y sustituirla por otra sustancia.

La ciencia había de considerar también las causas del cambio. Para Aristóteles había cuatro tipos de causas. La primera era la causa material o inmanente; para una estatua de bronce, el bronce es la causa inmanente. La segunda era la causa formal; para una estatua era el diseño o la forma. La causa formal de la armonía es la relación de 2 a 1 en la octava. La tercera causa era la causa eficiente, el agente o el actor; el artista y su cincel son las causas efectivas para la estatua. La cuarta

era la causa final o el propósito para el que servía el fenómeno; las estatuas sirven para el goce del pueblo, para ofrecer belleza. La causa final era la más importante de las cuatro porque daba la razón última de acontecimientos o fenómenos. Cada cosa tenía una causa final.

¿Qué lugar tenía la matemática en este esquema de cosas? Las ciencias físicas eran fundamentales para el estudio de la naturaleza y las matemáticas ayudaban a describir propiedades formales tales como la forma y la cantidad. Proporcionaban también explicaciones de hechos observados en fenómenos materiales. De esta manera, la geometría daba las razones de hechos que se producían en óptica y astronomía y la aritmética las proporciones que producirían la armonía. Pero la matemática era definitivamente una abstracción del mundo real, ya que los objetos matemáticos no son independientes o anteriores a la experiencia. Existen en la mente humana como una clase de ideas intermedias entre los objetos sensibles y la esencia de los mismos. Puesto que se han abstraído del mundo físico, son aplicables a él, pero no tienen ninguna realidad aparte de las cosas visibles y tangibles. La matemática sola no puede proporcionar nunca una definición adecuada de la sustancia. Diferencias cualitativas, como los distintos colores, no pueden ser reducidas a diferencias geométricas. En consecuencia, en el estudio de las causas, la matemática puede dar, en el mejor de los casos, algún conocimiento de la causa formal, esto es, una descripción. Puede describir lo que ocurre en el mundo físico, puede establecer correlaciones entre variaciones concomitantes, pero no puede decir nada acerca de las causas finales y efectivas del movimiento o el cambio. Así pues, Aristóteles distinguía formalmente entre matemática y física, y asignaba un papel menor a la matemática. No estuvo interesado en la predicción.

A partir de este resumen podemos ver que todos los filósofos griegos que forjaron y moldearon el mundo intelectual griego pusieron el énfasis en el estudio de la naturaleza para la comprensión y apreciación de su realidad subyacente. Desde los tiempos de los pitagóricos prácticamente todos aseguraban que la naturaleza estaba diseñada matemáticamente. Durante el período clásico, la teoría del diseño matemático de la naturaleza quedó establecida y la investigación de las leyes matemáticas, institucionalizada. A pesar de que esta teoría no motivaba todas la matemática creada después, una vez establecida fue aceptada y seguida concienzudamente por la mayoría de los grandes matemáticos. Durante el tiempo en que imperó esta doctrina, que fue hasta finales del siglo XIX, la investigación del diseño

matemático se identificó con la búsqueda de la verdad. Aunque algunos griegos —por ejemplo, Ptolomeo— sostenían que las teorías matemáticas eran solamente intentos humanos de proporcionar una descripción coherente, la creencia de que las leyes matemáticas eran la verdad en lo que se refiere a la naturaleza, atrajo hacia la matemática a algunos de los pensadores más nobles y profundos.

Observaremos también, con el fin de apreciar mejor lo que ocurrió en el siglo XVII, el énfasis griego sobre la potencia de la mente. Puesto que los filósofos griegos creían que la mente era el agente más poderoso para abarcar la naturaleza, adoptaron principios básicos que se referían a la mente. Así, la creencia de que el movimiento circular era el básico, defendida por Aristóteles sobre la base de que el círculo es completo mientras que una figura lineal, debido a que está limitada por varias curvas (segmentos lineales), es incompleta y por tanto de importancia secundaria, recurría a la mente basándose en principios estéticos. Que los cuerpos celestes se moverían sólo con velocidad constante o uniforme fue un concepto que agradaba a la mente, seguramente porque era más sencillo que el movimiento no uniforme. La combinación de los movimientos circular y uniforme pareció sentar bien a los cuerpos celestes. Que los cuerpos sublunares fueran diferentes de los planetas, el Sol y las estrellas también pareció razonable, ya que los cuerpos celestes presentaban un aspecto permanente mientras que los cambios sobre la Tierra eran evidentes. Incluso Aristóteles, que daba importancia a las abstracciones solamente en la medida en que podían servir de ayuda para comprender el mundo observable, decía que debemos comenzar a partir de principios que son conocidos y manifiestos a la mente y proceder entonces a analizar las cosas que se encuentran en la naturaleza. Vamos, dice, de lo universal a lo particular, del hombre a los hombres, exactamente como los niños llaman padre a todos los hombres y entonces aprenden a distinguir. Así pues, incluso las abstracciones que vienen de objetos concretos presuponen algunos principios generales que emanan de la mente. Esta doctrina, el poder de la mente para producir primeros principios, fue rechazada el siglo XVII.

4. La astronomía matemática griega

Vamos a examinar ahora la producción griega en lo que se refiere a la descripción matemática de los fenómenos naturales. Es a partir de

los tiempos de Platón cuando las distintas ciencias creadas por los griegos toman un contenido y una dirección significativos. Aun cuando nos proponemos analizar la astronomía griega, lo haremos a través de un aspecto de la geometría euclídea. Ya hemos observado que la geometría esférica se desarrolló en función de la astronomía. La geometría era, en efecto, subsidiaria de un campo más general como es la cosmología. Los principios geométricos, para los griegos, estaban incorporados a la estructura global del Universo, de la que el espacio era el principal componente. Por tanto, el estudio del propio espacio y de las figuras del mismo era de importancia para un fin superior. En otras palabras, la geometría fue una ciencia: la ciencia del espacio físico.

Fue Platón quien, pese a estar enterado del número impresionante de observaciones astronómicas hechas por los babilonios y los egipcios, puso énfasis en la ausencia de una teoría subyacente o unificadora y de una explicación de los movimientos aparentemente irregulares de los planetas. Eudoxo, que fue durante un tiempo alumno de la Academia, hizo suyo el problema de Platón de «salvar las apariencias». Su respuesta es la primera teoría astronómica razonablemente completa. Escribió cuatro libros sobre astronomía: *Espejo, Acontecimientos, El periodo de ocho años y Sobre velocidades,* de los que conocemos sólo algunos fragmentos. Por estos fragmentos y los relatos de otros autores conocemos el espíritu de la teoría de Eudoxo.

Los movimientos del Sol y la Luna, vistos desde la Tierra, pueden ser descritos a grandes rasgos como circulares con velocidad constante. Sin embargo, sus desviaciones de las órbitas circulares son lo suficientemente grandes como para haber sido observadas y para requerir una explicación. Los movimientos de los planetas, observados desde la Tierra, son todavía más complejos, pues durante una revolución cualquiera invierten su recorrido, marchan hacia atrás durante un tiempo y de nuevo vuelven a moverse hacia adelante. Además, sus velocidades en estos caminos son variables.

Para demostrar que los movimientos reales, bastante complicados y aparentemente sin leyes, eran comprensibles en términos de movimientos circulares sencillos, Eudoxo propuso el esquema siguiente: para cualquier cuerpo celeste existía un conjunto de tres o cuatro esferas, todas concéntricas y cuyo centro era la Tierra, girando cada una de ellas alrededor de un eje. La esfera más interior contenía el cuerpo que se movía a lo largo de lo que se podría llamar el ecuador de dicha esfera; es decir, el eje de rotación era perpendicular al camino

circular del cuerpo. Sin embargo, mientras giraba sobre su eje, esta esfera era arrastrada por la rotación de la esfera concéntrica siguiente por medio del siguiente artificio: imaginemos que el eje de rotación de la primera esfera se prolonga por cada uno de sus extremos hasta cortar la segunda esfera. Si ahora la segunda esfera gira alrededor de su propio eje, hará girar también el eje de la primera al mismo tiempo que esta última da vueltas alrededor del mismo. El eje de la segunda esfera es a su vez arrastrado por la rotación de la tercera esfera sobre su eje. Eudoxo encontró que para el Sol y la Luna bastaba una combinación de tres esferas para reproducir sus movimientos reales observados desde la Tierra. Para cada planeta se requería una cuarta esfera, relacionada con la tercera como ya hemos descrito. La esfera más exterior daba una vuelta cada 24 horas alrededor de un eje que pasaba por los polos de la esfera celeste. En total, Eudoxo usaba 27 esferas. Sus ejes de rotación, velocidades de giro y radios estaban determinados de manera que la teoría se ajustara lo mejor posible a las observaciones disponibles en su tiempo.

El esquema de Eudoxo era matemáticamente elegante y destacable en varios aspectos. La idea misma de la utilización de combinaciones de esferas era ingeniosa; y la tarea de elegir los ejes, velocidades y radios para hacer que el movimiento resultante del cuerpo celeste se ajustara a las observaciones reales requiere grandes habilidades matemáticas para trabajar con superficies y curvas (esto es, los caminos de los planetas) en el espacio.

Es especialmente digno de mención que la teoría de Eudoxo es estrictamente matemática. Sus esferas, excepto la «esfera» de las estrellas fijas, no eran cuerpos materiales sino construcciones matemáticas. Tampoco intentaba describir las fuerzas que harían girar las esferas tal como él decía que lo hacían. Su teoría es completamente moderna en su espíritu, pues hoy en día la finalidad de la ciencia consiste en dar una descripción matemática de los fenómenos y no una explicación física de los mismos.

El sistema de Eudoxo tenía serios inconvenientes. No tenía en cuenta la velocidad variable del Sol y presentaba pequeños errores acerca de su camino real. Su teoría no acababa de encajar con el movimiento real de Marte y tampoco era satisfactoria para Venus. Que Eudoxo admitiera tales inconvenientes puede explicarse por el hecho de que creía que no tenía en su poder el número de observaciones suficiente. Probablemente había estudiado en Egipto solamente los hechos más importantes acerca de puntos estacionarios, regresio-

nes y períodos de revolución de los planetas exteriores (Marte, Júpiter y Saturno). Quizá también por este motivo sus valores para los tamaños y las distancias de los cuerpos celestes estaban muy equivocados. Aristarco dice que Eudoxo creía que el diámetro del Sol era nueve veces el de la Luna.

Aristóteles no valoraba un esquema estrictamente matemático, por lo que no le satisfizo la solución de Eudoxo. Para inventar un mecanismo real que haga que una esfera obligue a otra a girar, añadió 29 esferas. Estas estaban intercaladas entre las de Eudoxo de tal forma que una esfera arrastra a otra a través de un contacto real y por ello todas transmitían su propia potencia a partir de la más exterior. En algunos escritos de Aristóteles la esfera de las estrellas, que era móvil, es el primer motor de las restantes. En otros existe un motor inmóvil detrás de la esfera de estrellas. Sus 56 esferas complicaron el sistema de tal manera que fue desacreditado por los científicos, pese a que fue popular entre seglares en la Europa medieval. Aristóteles también creía que la Tierra era esférica, por razones de simetría y equilibrio y porque la sombra de la Tierra sobre la Luna, visible en los eclipses lunares, es circular.

Los escritos sobre astronomía continúan después de Aristóteles sin solución de continuidad. Después de los trabajos de Autólico (cap. 3, sec. 10) y los *Fenómenos* de Euclides (cap. 4, sec. 11) los siguientes grandes trabajos astronómicos son alejandrinos. Las observaciones de los caldeos, las observaciones hechas por los babilonios del período seléucida y las mediciones hechas por los propios alejandrinos incrementaron enormemente el número y la precisión de los datos asequibles.

El primer gran astrónomo alejandrino fue Aristarco (sobre 310-230 a. C.) que estudió geometría, astronomía, música y otras ramas de la ciencia. Su libro *Sobre el Tamaño y las Distancias del Sol y la Luna*, del que existen manuscritos en griego y árabe, es el primer gran intento de medir las distancias del Sol y la Luna desde la Tierra y los tamaños relativos entre estos cuerpos. Estos cálculos son, claramente, otro ejemplo del interés de los alejandrinos respecto de cuestiones de tipo cuantitativo. Aristarco no tenía a su disposición la trigonometría, ni un buen valor para π (el trabajo de Arquímedes al respecto apareció más tarde), pero utilizó la geometría euclídea de forma muy eficaz.

Sabía que la luz de la Luna es luz reflejada. Cuando exactamente la mitad de la Luna está iluminada, el ángulo en M (fig. 7.1) es un ángulo recto. El observador en O puede medir el ángulo allí y, entonces, dar

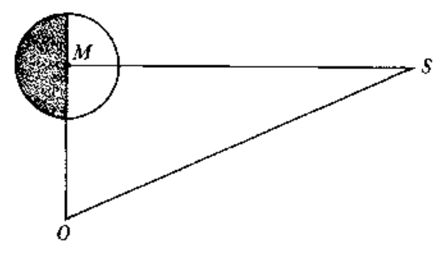


FIGURA 7.1

una estimación de las distancias *relativas* OM y OS. La medición del ángulo de Aristarco fue de 87°; aunque el valor aproximado es de 89°52′. En consecuencia, estimó que el Sol está más de 18 veces y menos de 20 más alejado que la Luna. La relación correcta es 346 veces más distante.

Una vez que conoce las distancias relativas, Aristarco calcula los tamaños relativos midiendo los tamaños de los discos del Sol y de la Luna visibles desde la Tierra. Concluyó que el volumen del Sol era 7000 veces mayor que el de la Luna. Estuvo muy lejos de la verdad: el número correcto es 64.000.000. Calculó también que la razón entre el radio del Sol y el de la Tierra está comprendida entre 19/3 y 43/6; no obstante, la razón correcta es alrededor de 107.

Aristarco es igualmente famoso por haber sido el primero que propuso la hipótesis heliocéntrica —la Tierra y los planetas dan vueltas alrededor del Sol, que permanece fijo, describiendo círculos—. Las estrellas también están fijas y su movimiento aparente se debe a la rotación de la Tierra sobre su eje. La Luna gira alrededor de la Tierra. Pese a que, como sabemos hoy día, Aristarco estaba en lo cierto, su idea no fue aceptada por varias razones. Por una parte, la mecánica griega (ver más abajo), que había sido ya bien desarrollada por Aristóteles, no consideraba objetos situados en una Tierra móvil. De acuerdo con Aristóteles, los objetos pesados buscaban el centro del Universo. Este principio daba cuenta de la caída de objetos hacia la Tierra, ya que ésta era el centro del Universo; pero si estuviera en movimiento los objetos quedarían detrás de ella. Este argumento fue utilizado por Ptolomeo contra Aristarco y, de hecho, fue usado también contra Copérnico ya que el sistema mecánico que seguía prevaleciendo era todavía el de Aristóteles. Ptolomeo decía también que las nubes se rezagarían detrás de una Tierra en movimiento. Además, la Mecánica de Aristóteles requería una fuerza para mante-

ner los objetos terrestres en movimiento y no existía ninguna fuerza evidente. No conocemos la forma en que Aristarco rebatió estos argumentos.

Otro argumento que se lanzó contra Aristarco era que si la Tierra estuviera en movimiento su distancia a las estrellas fijas debía variar, pero esto claramente no era así. Aristarco dio a esto la refutación correcta: decía que el radio de la esfera de las estrellas fijas es tan grande que la órbita de la Tierra es demasiado pequeña para apreciar tal variación. La idea heliocéntrica de Aristarco fue rechazada por muchos por considerarla impía, al identificar la materia corruptible de la Tierra con la materia incorruptible de los cuerpos celestes. La hipótesis de que los planetas se mueven alrededor del Sol describiendo círculos es, naturalmente, insatisfactoria, ya que el movimiento es realmente más complicado, pero la idea heliocéntrica podía haber sido refinada y Copérnico lo hizo más tarde. No obstante, fue demasiado radical para el pensamiento griego.

El creador de la Astronomía matemática cuantitativa es Apolonio. Era llamado Epsilon debido a que el símbolo ε se usaba para designar la Luna y gran parte de su astronomía estaba dedicada al movimiento de dicho astro. Antes de considerar su trabajo así como el de Hiparco y Ptolomeo, con los que está fuertemente relacionado, examinaremos el esquema básico que había inspirado la astronomía griega entre los tiempos de Eudoxo y Apolonio: el esquema de epiciclo y deferente. En este esquema un planeta P se mueve con velocidad constante sobre un círculo (fig. 7.2) de centro S, mientras que S se mueve a su vez con velocidad constante sobre un círculo centrado en la Tierra E. El círculo sobre el que se mueve S recibe el nombre de deferente; el círculo sobre el que se mueve P, epiciclo. Para algunos planetas, el punto S es el Sol, pero en otros casos es solamente un punto

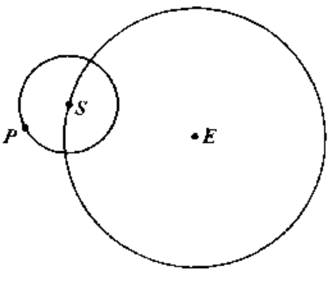


FIGURA 7.2

matemático. El sentido del movimiento de P puede coincidir o ser opuesto al del movimiento de S. Este último es el caso del Sol y de la Luna.

Se supone que Apolonio conocía perfectamente el esquema del movimiento epicíclico y los detalles que permiten la representación de los movimientos de los planetas, la Luna y el Sol. Ptolomeo atribuye específicamente a Apolonio la determinación de los puntos en que un planeta permanece estacionario e invierte el sentido de su movimiento.

El punto culminante de la astronomía griega son los trabajos de Hiparco y Ptolomeo. El esquema de deferente y epiciclo fue tomado por Hiparco (muerto sobre el 125 a. C.) y aplicado al movimiento de los cinco planetas conocidos entonces, la Luna, el Sol y las estrellas. Conocemos los trabajos de Hiparco a través del *Almagesto* de Ptolomeo, si bien resulta difícil distinguir qué es debido a Hiparco y qué a Ptolomeo. Después de realizar observaciones en Rodas durante treinta y cinco años y utilizar las observaciones de los babilonios, Hiparco trabajó sobre los detalles de la teoría epicíclica del movimiento. Seleccionando adecuadamente los radios del epiciclo y el deferente y las velocidades de un cuerpo sobre su epiciclo así como el centro de éste sobre el deferente, era capaz de obtener una descripción mejorada de los movimientos. Tuvo mucho éxito con el Sol y la Luna, pero su éxito fue solamente parcial con los planetas. Durante el tiempo de Hiparco, un eclipse de Luna se predecía con un error de una o dos horas, mientras que los eclipses de Sol se predecían con menor precisión. Esta teoría era válida también para las estaciones.

La aportación más original de Hiparco es su descubrimiento de la precesión de los equinoccios. Para comprender este fenómeno supongamos que el eje de rotación de la Tierra se prolonga hasta las estrellas. El punto en el que corta la esfera de las estrellas describe un círculo y tarda 2600 años en recorrerlo. En otras palabras, el eje de la Tierra varía continuamente su dirección respecto a las estrellas y el movimiento es periódico. La estrella a la que apunta en cualquier momento recibe el nombre de Estrella Polar. El ángulo abarcado desde la Tierra por su diámetro del círculo mencionado con anterioridad es de 45°.

Hiparco tiene varias contribuciones más a la astronomía, como la construcción de instrumentos para la observación, la determinación del ángulo de la eclíptica, la medida de irregularidades en el movimiento de la Luna, la mejora de la determinación de la duración del año solar (que estimó en 365 días, 5 horas, 55 minutos y 12 segundos

(aproximadamente 6 1/2 minutos demasiado largo) y un catálogo de alrededor de mil estrellas. Encontró que la razón entre la distancia a la Luna y el radio de la Tierra era 67,74; el verdadero valor es 60,4. Calculó que el radio de la Luna es 1/3 del de la Tierra; sabemos en la actualidad que es 27/100.

Ptolomeo extendió los trabajos de Hiparco con varias mejoras sobre la descripción matemática de los movimientos de todos los cuerpos celestes. La teoría geocéntrica generalizada, presentada en el *Almagesto*, ofrece una exposición completa de la teoría de deferente y epiciclo, que se conoce hoy en día como teoría de Ptolomeo.

Con el fin de dar una descripción geométrica de las observaciones realizadas, Ptolomeo introdujo también una modificación del movimiento epicíclico conocido como el movimiento ecuante uniforme. En este esquema (fig. 7.3) el planeta se mueve sobre un epiciclo con centro en Q al tiempo que Q se mueve sobre un círculo de centro C, que no es la Tierra sino un punto próximo. Fija la velocidad de Q introduciendo el punto R tal que $EC = CR$ y de manera que $\measuredangle QRT$ crece uniformemente. Así Q se mueve con velocidad angular constante, pero no con velocidad lineal uniforme.

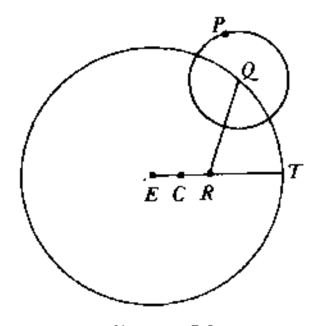


FIGURA 7.3

El método y conocimientos que alcanzaron los astrónomos griegos son completamente modernos. Los astrónomos greco-alejandrinos, en especial Hiparco y Ptolomeo, hicieron sus propias observaciones; de hecho, Hiparco no confiaba en muchas de las antiguas observaciones egipcias y caldeas y las repitió de nuevo. Los astrónomos clásicos y los alejandrinos no sólo construyeron teorías sino que comprobaron completamente que estas teorías no constituían el verdadero diseño sino descripciones que encajaban con las

observaciones. Ptolomeo dice en el *Almagesto* [1] que en astronomía era necesario buscar un modelo matemático lo más sencillo posible. Estos sabios, al contrario que otros griegos, no buscaban una explicación física de los movimientos. Acerca de esto, Ptolomeo dice [2]: «Después de todo, hablando con generalidad, la causa de los primeros principios es o bien nada o bien algo difícil de interpretar en su naturaleza.» Pero su propio modelo matemático fue tomado más tarde como la verdad en sentido literal por el mundo cristiano.

La teoría ptolemaica ofreció la primera evidencia razonablemente completa de la uniformidad e invariabilidad de la Naturaleza y fue la última respuesta griega al problema de Platón de racionalizar los movimientos aparentes de los cuerpos celestes. Ninguna otra producción de toda Grecia podía rivalizar con el *Almagesto* debido a su profunda influencia sobre las concepciones del Universo y ninguna, salvo los *Elementos* de Euclides, logró tan incuestionable autoridad.

Este breve relato de la astronomía griega no ha revelado todo el fondo y extensión del trabajo realizado incluso por los autores citados aquí, y omite muchas contribuciones más. Casi todos los matemáticos griegos, incluido Arquímedes, se dedicaron a esta cuestión. La astronomía griega fue dominante y globalizadora y utilizó muchas matemáticas.

5. La Geografía

Otra ciencia creada en tiempos de los griegos es la Geografía. Aunque algunos griegos clásicos como Anaximandro y Hecateo de Mileto (muerto sobre el 475 a. C.) hicieron algunos mapas de la Tierra tal como era conocida entonces, fueron los alejandrinos quienes realizaron los grandes avances en geografía. Midieron o calcularon distancias a lo largo de la Tierra, la altura de las montañas, la profundidad de los valles y la extensión de los mares. Los alejandrinos estaban especialmente estimulados a estudiar geografía porque el mundo griego se había extendido.

El primer gran geógrafo alejandrino fue Eratóstenes de Cirene (*c.* 284-*c.* 192 a. C.), director de la biblioteca de Alejandría, matemático,

---

[1] Libro XIII, cap. 2, último parágrafo.

[2] *Almagesto,* Libro IX.

poeta, filósofo, historiador, filólogo, cronólogo y con la fama de ser uno de los hombres más cultos del mundo antiguo. Estudió en Atenas en la escuela de Platón y fue invitado a Alejandría por Ptolomeo Euergetes. Eratóstenes trabajó en Alejandría hasta que a su vejez le sobrevino la ceguera; a causa de esta desgracia dejó de comer hasta morir de ese modo.

Eratóstenes recopiló todo el conocimiento geográfico del momento y realizó numerosos cálculos de distancias sobre la Tierra entre lugares importantes (como por ejemplo las ciudades). Uno de sus cálculos más famosos es la longitud de la circunferencia de la Tierra. Al mediodía del solsticio de verano, observó que el Sol estaba prácticamente sobre la vertical de Siena, la ciudad llamada actualmente Asuán (fig. 7.4). (Esto fue confirmado al observar que el Sol brillaba directamente vertical allí hasta el fondo de un pozo.) Al mismo tiempo, en Alejandría, que está (con una variación de 1°) en el mismo meridiano que Siena pero situada más al Norte, observó que el ángulo entre la vertical (*OB* en la figura) y la dirección del Sol (*AD* en la figura) era de 1/50 de 360°. El Sol está suficientemente alejado de la Tierra, por lo que podemos considerar que *SE* y *AD* son paralelas. Luego $\sphericalangle SOA$ es 1/50 × 360°. Esto significa que el arco *SA* es 1/50 de la circunferencia de la Tierra. Eratóstenes dio una estimación de la distancia entre Alejandría y Siena usando el hecho de que un convoy de camellos, que por lo común viajaba 100 estadios diarios, tardaba 50 días en llegar a Siena. Por lo tanto, esta distancia es de 5000 estadios y la circunferencia de la Tierra mide 250.000 estadios. Se cree que un estadio equivalía a 157 metros por lo que el resultado de Eratóstenes es de 24.662 millas. Este resultado era mucho más exacto que todos los resultados anteriores.

Eratóstenes escribió una *Geografía* en la que incorporó los métodos y resultados de sus mediciones y cálculos. Incluye explicaciones acerca de la naturaleza y las causas de los cambios que han tenido lugar sobre la superficie de la Tierra. Hizo también un mapa del mundo.

El trazado científico de mapas se convirtió en una parte de la geografía. Está generalmente admitido que fue Hiparco quien introdujo la latitud y la longitud, si bien el sistema era conocido con anterioridad. El uso de la latitud y la longitud permite una descripción precisa de localizaciones sobre la Tierra. Hiparco inventó la proyección ortográfica, por la cual los «rayos de luz» procedentes del infinito proyectan la Tierra sobre un plano (fig. 7.5). Nuestra visión

de la Luna es prácticamente ortográfica. Este método le permitió representar una porción de la Tierra sobre una superficie plana.

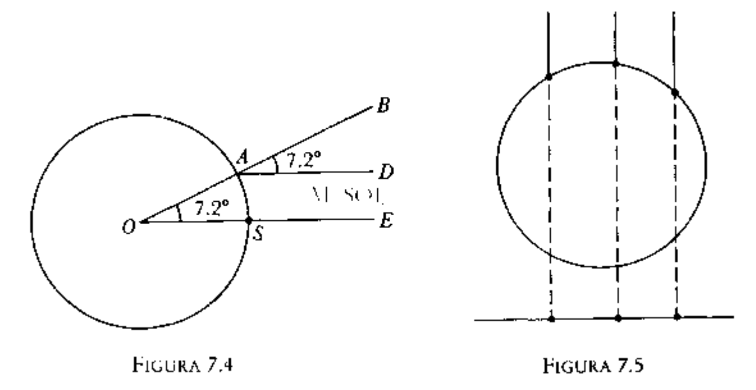


FIGURA 7.4 FIGURA 7.5

Ptolomeo en su *Planisferio*, y probablemente Hiparco antes que él, empleó el método de la proyección estereográfica. Una línea que parte de O (fig. 7.6) y pasa por *P*, situado sobre superficie de la Tierra, se prolonga hasta que corte el plano del Ecuador o un plano tangente por el polo opuesto. Hiparco usaba presumiblemente este último y Ptolomeo el plano ecuatorial. De esta manera los puntos de la esfera se transfieren a un plano. En este esquema todos los puntos del mapa están en la verdadera dirección desde el centro del mismo. Asimismo, los ángulos quedan conservados localmente (transformación conforme), aunque Ptolomeo no menciona nada de esto. Los meridianos y los paralelos de las latitudes son perpendiculares. Los círculos de la esfera se convierten en círculos del plano, pero el área no se conserva. El propio Ptolomeo ideó la proyección cónica, es decir, la proyección de una región de la Tierra desde el centro de la esfera sobre un cono tangente (fig. 7.7).

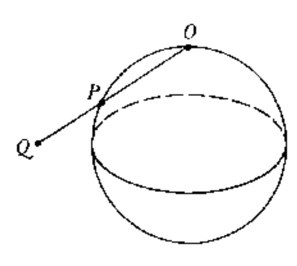


FIGURA 7.6

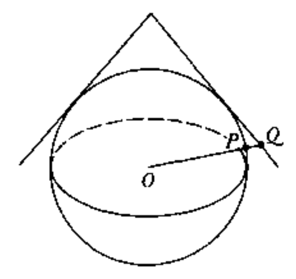


FIGURA 7.7

En su *Geografía*, una obra en ocho tomos, Ptolomeo enseña métodos de confección de mapas. El capítulo 24 del libro I es el trabajo más antiguo que existe dedicado por su título y su contenido a aplicar una esfera sobre un plano. La *Geografía* completa es el primer atlas y diccionario geográfico. Da la latitud y la longitud de 8.000 lugares de la Tierra y fue la referencia típica durante cientos de años.

6. La Mecánica

Los griegos fueron los iniciadores de la ciencia de la Mecánica. En su *Física*, Aristóteles incluye una teoría del movimiento que es el punto culminante de la mecánica griega. Como toda su física, su mecánica está basada en principios racionales, aparentemente semievidentes, sólo levemente comprobados por la observación y la experiencia o diseñados a partir de ellas.

Según Aristóteles, hay dos tipos de movimiento, el natural y el violento o creado por el hombre. Las esferas celestes tienen solamente un movimiento natural, que es circular. Para los objetos terrestres, él pensaba que poseen un movimiento natural (en oposición a los movimientos violentos provocados al arrastrar o empujar un cuerpo de un lugar a otro) debido a que cada cuerpo tiene un lugar natural en el Universo en el que permanece en equilibrio o en reposo. Los cuerpos pesados tienen su lugar natural en el centro del Universo, que coincide con el centro de la Tierra. Los movimientos naturales se producen cuando un cuerpo busca su lugar natural. En sus movimientos naturales, los cuerpos terrestres describen trayectorias rectas hacia arriba o hacia abajo. Si un objeto terrestre no estuviera en su lugar natural, lo buscaría con la mayor rapidez posible. Los movimientos violentos, es decir, los provocados por el hombre, están compuestos de partes circulares y de partes rectilíneas. Así, una piedra lanzada hacia arriba sigue un camino en línea recta hacia arriba y un camino en línea recta hacia abajo.

Cualquier cuerpo en movimiento está sujeto a una fuerza y a una resistencia, la fuerza es el peso del cuerpo y la resistencia viene del medio en el que se mueve dicho cuerpo. En el movimiento violento la fuerza está aplicada por la mano o por algún mecanismo y la resistencia procede de su peso. Sin fuerza no habría movimiento; sin resistencia el movimiento quedaría completado en un instante. La velocidad de cualquier movimiento, entonces, depende de la fuerza y de la

resistencia. Estos principios se pueden resumir de forma moderna por la fórmula $V \propto F/R$; esto es, la velocidad es directamente proporcional a la fuerza e inversamente proporcional a la resistencia.

Como en el movimiento violento la resistencia está provocada por el peso, para cuerpos ligeros la resistencia, R, es menor. En virtud de la fórmula anterior, la velocidad debe ser mayor; es decir, los cuerpos más ligeros se mueven con más rapidez bajo la misma fuerza. En el caso del movimiento natural, la fuerza es el peso, por lo que cuerpos más pesados caen más rápidamente. Como en un movimiento natural la resistencia viene dada por el medio, en el vacío la velocidad debería ser infinita. Por lo tanto el vacío es imposible.

Aristóteles tenía dificultades para describir algunos fenómenos. Para explicar las velocidades crecientes en la caída de los cuerpos hacía que la velocidad del cuerpo aumentara a medida que se iba aproximando a su lugar natural porque el cuerpo se mueve con más alegría; pero esto no es coherente con que la velocidad dependa del peso fijo. En el caso de una flecha lanzada desde un arco, Aristóteles decía que la flecha continuaba el movimiento aunque no estuviera en contacto con el arco porque la mano o el arquero comunicaba una potencia de movimiento al aire circundante y este aire a la capa siguiente de aire y así sucesivamente. Alternativamente, el aire situado frente a la flecha se comprimía y empujaba alrededor de la parte posterior de la flecha para prevenir un vacío y así la flecha es empujada hacia adelante. No explicó la caída de la potencia motriz.

El mayor físico matemático de los tiempos de Grecia es Arquímedes. El, más que ningún otro autor griego, acercó la geometría a la mecánica y utilizó con gran ingenio argumentos geométricos para dar sus demostraciones. En mecánica escribió *Sobre el equilibrio de planos o el centro de gravedad de planos,* una obra en dos tomos. Por centro de gravedad de un cuerpo o una colección de cuerpos enlazados rígidamente entre sí entiende, igual que lo hacemos nosotros, el punto en el que se debe apoyar el cuerpo o colección de cuerpos para que esté en equilibrio por la fuerza de la gravedad. Comienza con postulados sobre la palanca y el centro de gravedad. Por ejemplo (los números están de acuerdo con los asignados por Arquímedes):

1. Pesos iguales a distancias iguales están en equilibrio y pesos iguales a distancias distintas no están en equilibrio, pero el peso se inclina hacia el que está a mayor distancia.

2. Si, cuando pesos situados a determinadas distancias están en equilibrio, se añade algo a uno de ellos, dejan de estar en equilibrio pero se inclinan hacia el peso al que se ha hecho la adición.

5. En figuras distintas pero semejantes sus centros de gravedad estarán situados de manera que guardarán la misma semejanza...

7. En cualquier figura cuyo perímetro es cóncavo en la misma dirección el centro de gravedad debe estar en el interior de la figura.

A continuación de estos postulados coloca un número determinado de proposiciones; las demostraciones de algunas de ellas dependen de resultados de un tratado perdido, *Sobre Palancas:*

Proposición 4. Si dos pesos iguales no tienen el mismo centro de gravedad, el centro de gravedad del sistema formado por los dos cuerpos está en el punto medio del segmento cuyos extremos son los centros de gravedad respectivos.

Proposiciones 6 y 7. Dos magnitudes tanto conmensurables como inconmensurables se equilibran en una balanza a distancias inversamente proporcionales a las magnitudes.

Proposición 10. El centro de gravedad de un paralelogramo es el punto de intersección de sus diagonales.

Proposición 14. El centro de gravedad de un triángulo está en el punto de intersección de las líneas trazadas desde dos vértices cualesquiera al punto medio del lado opuesto respectivo.

El libro II se dedica al centro de gravedad de un segmento parabólico. Entre los teoremas principales están:

Proposición 4. El centro de gravedad de un segmento parabólico determinado por una línea está situado sobre el diámetro del segmento.

El diámetro es AO (fig. 7.8), donde O es el punto medio de BD y AO es paralelo al eje de la parábola. La demostración utiliza resultados obtenidos en su *Cuadratura de la parábola.*

Proposición 8. Si AO es el diámetro de un segmento parabólico y G su centro de gravedad, entonces $AG = (3/2)GO$.

Trabajos sobre centros de gravedad pueden encontrarse en varios libros del período greco-alejandrino. Algunos ejemplos son la *Mecá-*

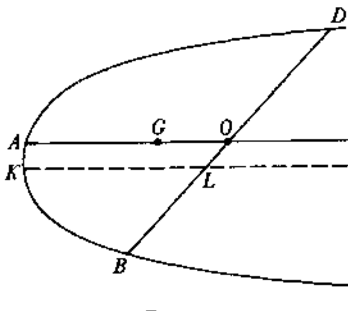


FIGURA 7.8

nica de Herón y el libro VIII de la *Colección Matemática* de Pappus (cap. 5, sec. 7).

La ciencia de la hidrostática —el estudio de la presión de los fluidos en reposo— fue creada por Arquímedes. En su libro *Sobre los cuerpos flotantes* están interesado en la presión ejercida por el agua sobre los objetos situados en ella. Da dos postulados. El primero de ellos es que el efecto que produce la presión ejercida por cualquier parte del fluido sobre el fluido es descendente. El segundo postulado afirma que la presión ejercida por un fluido sobre un cuerpo situado en él está dirigida hacia arriba en la dirección de la perpendicular que pasa por el centro de gravedad del cuerpo. Algunos de los teoremas que demuestra en el libro I son:

Proposición 2. La superficie de un fluido en reposo es la superficie de una esfera cuyo centro es el centro de la Tierra.

Proposición 3. Los sólidos que, a tamaños iguales, tienen el mismo peso que un fluido, si se hunden en un fluido quedarán sumergidos de manera que no emergen por encima de la superficie, pero tampoco se hunden por debajo de ella.

Proposición 5. Si se sitúa en un fluido un sólido más ligero que el propio fluido quedará sumergido en él de manera que el peso del sólido en el aire será igual al peso del fluido desplazado.

Proposición 7. Un sólido más pesado que un fluido desciende hasta el fondo del mismo cuando se sitúa en él, y el sólido, cuando se pese en el fluido, será más ligero que su verdadero peso en el peso del fluido desplazado.

Se cree que esta última proposición es una de las empleadas por Arquímedes para determinar la composición de una famosa corona (cap. 5, sec. 3). Debió razonar de la manera siguiente: sea W el peso de la corona. Tomemos una corona de oro puro de peso W y pesémosla en el fluido. Pesará una cantidad F_1 menos, que es el peso del agua desplazada. Análogamente, un peso W de plata pura desplazará un peso de agua igual a F_2, que puede calcularse pesando la plata en el agua. Entonces, si la corona original contiene un peso w_1 de oro y un peso w_2 de plata, la corona original debe desplazar un peso de agua igual a

$$\frac{w_1}{W}F_1 + \frac{w_2}{W}F_2.$$

Sea F el peso real del agua desplazada por la corona. Entonces

$$\frac{w_1}{W}F_1 + \frac{w_2}{W}F_2 = F$$

o

$$w_1F_1 + w_2F_2 = (w_1 + w_2)F,$$

o bien

$$\frac{w_1}{w_2} = \frac{F_2 - F}{F - F_1}.$$

De esta manera Arquímedes fue capaz de determinar la razón entre las cantidades de oro y de plata que contenía la corona. La narración de Vitrubio de esta historia relata que Arquímedes uso *volúmenes* de agua desplazada en vez de pesos. En este caso, los valores F, F_1 y F_2 anteriores son, respectivamente, los volúmenes de agua desplazada por la corona, un peso W de oro puro y un peso W de plata pura. Se utilizan las mismas fórmulas, pero ahora no se debe utilizar la proposición 7. Arquímedes halló que el oro había sido degradado con plata.

Para apreciar un poco las complejidades matemáticas y físicas de los problemas tratados por Arquímedes en este trabajo, citaremos una de las proposiciones sencillas del libro II.

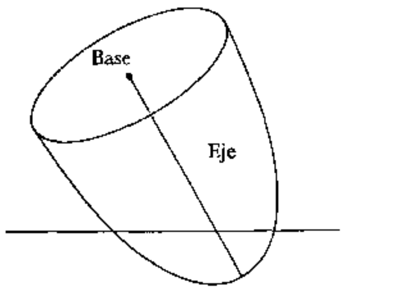
Proposición 2. Si un segmento recto de un paraboloide de revolución cuyo eje no es mayor que $3p/4$ [p es el parámetro principal o *«latus rectum»* de la parábola generatriz] y cuya gravedad específica es menor que la de un fluido, se sumerge en el fluido con su eje inclinado un determinado ángulo respecto de la vertical, pero de manera que la base del segmento no toque la superficie del fluido (fig. 7.9), el segmento del paraboloide no permanecerá en esta posición sino que volverá a la posición en la que el eje esté vertical.

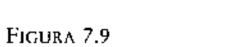
El tema que trata Arquímedes es la estabilidad de cuerpos situados en el agua. Demuestra bajo qué condiciones un cuerpo situado en el agua se dará la vuelta o permanecerá en equilibrio. Los problemas son evidentemente idealizaciones de cómo se comportarían los barcos cuando estuvieran obligados a tomar diferentes inclinaciones en el agua.

7. La Optica

Después de la Astronomía, la Optica ha sido una de las ciencias matemáticas más cultivadas y de más éxito. Fue creada por los griegos. Casi todos los filósofos griegos, comenzando por los pitagóricos, especulaban acerca de la naturaleza de la luz, la visión y el color. Nosotros nos fijaremos, sin embargo, en los logros matemáticos. El primero de ellos es la afirmación sobre los fundamentos a priori dada por Empédocles de Agrigento en Sicilia (*c.* 490 a. C.) de que la luz viaja con velocidad finita.

Los primeros tratamientos sistemáticos que poseemos son la *Optica* y la *Catóptrica*. La *Optica* se refiere al problema de la visión y al uso de la misma para la determinación de los tamaños de los objetos. Euclides comienza con definiciones (que en realidad son postulados), el primero de los cuales establece (como hacía Platón) que la visión es posible porque los rayos de luz emitidos por el ojo viajan a lo largo de líneas rectas e inciden sobre los objetos vistos. La definición 2 afirma que la figura formada por los rayos visuales es un cono cuyo vértice está situado en el ojo y su base, en los extremos del objeto. La definición 4 dice que de dos objetos, aparece como mayor el que determina un mayor ángulo en el vértice. Luego, en la proposición 8, Euclides demuestra que los tamaños aparentes de dos objetos iguales y paralelos (AB y CD en la fig. 7.10) no son proporcionales a sus distancias al ojo. Las proposiciones 23 a 27 prueban que




FIGURA 7.9

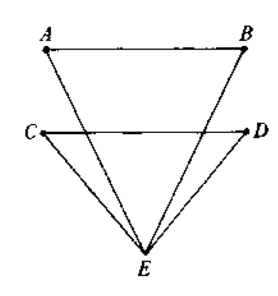


FIGURA 7.10

un ojo que mira una esfera, en realidad ve menos de la mitad de la misma y que el contorno de lo que ve es un círculo. Las proposiciones 32 a 37 aseguran que el ojo que mira un círculo verá un círculo solamente si el ojo está situado sobre la perpendicular al plano del círculo trazada por su centro. Euclides muestra también cómo calcular los tamaños de objetos vistos a través de un espejo plano. Hay 58 proposiciones en el libro.

La *Catóptrica* (teoría de los espejos) describe el comportamiento de los rayos de luz reflejados sobre espejos planos, cóncavos y convexos y el efecto de este comportamiento sobre lo que vemos. Igual que en la *Optica*, comienza con definiciones que realmente son postulados. El teorema 1, la ley de la reflexión, es ahora fundamental en lo que se llama óptica geométrica. Dice que el ángulo A, que es el formado por el rayo incidente con el espejo (fig. 7.11) es igual al ángulo B, que es el formado por el rayo reflejado con el espejo. Es más corriente hoy en día decir que $\sphericalangle C = \sphericalangle D$ y hablar de $\sphericalangle C$ como el ángulo de incidencia y $\sphericalangle D$ como del ángulo de reflexión. Euclides prueba también la ley para un rayo que incide sobre un espejo

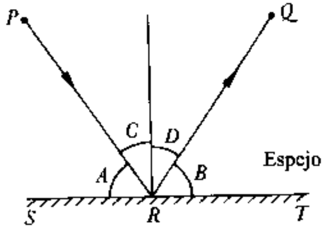


FIGURA 7.11

convexo o uno cóncavo sustituyendo el espejo por una tangente en el punto de incidencia del rayo.

Herón sacó una consecuencia importante de la ley de la reflexión. Si *P* y *Q* (fig. 7.11) son dos puntos cualesquiera situados a un mismo lado de la recta *ST*, de todos los caminos que podrían seguirse para ir del punto *P* a la línea y de aquí al punto *Q*, el más corto es el que pasa por el punto *R* de manera que las dos líneas *PR* y *QR* forman ángulos iguales con la recta —que es exactamente el camino que sigue un rayo de luz. Así pues, la luz sigue el camino más corto que va de *P* a *Q* a través del espejo. Aparentemente, la naturaleza conoce muy bien la geometría y la utiliza a su total conveniencia. Esta proposición aparece en la *Catóptrica* de Herón, que trata también de espejos cóncavos y convexos y combinaciones de espejos.

Se escribió una cantidad enorme de trabajos sobre la reflexión de la luz mediante espejos de varias formas. Entre ellos están la actualmente perdida *Catóptrica* de Arquímedes y los dos trabajos, llamados ambos *Sobre los espejos ustorios*, de Diocles y Apolonio. Los espejos incendiarios eran sin duda espejos cóncavos en forma de esferas, paraboloides de revolución y elipsoides, este último formado al girar una elipse alrededor de su eje mayor. Indudablemente, Apolonio conocía que un espejo en forma de paraboloide reflejaría la luz procedente de un foco a través de un haz paralelo al eje del espejo. Recíprocamente, los rayos que vienen paralelos al eje se concentrarán en el foco. Los rayos del Sol así concentrados producen un calor muy grande en el foco; de ahí el término de espejo incendiario. Esta es la propiedad del espejo en forma de paraboloide que se supone utilizó Arquímedes para concentrar los rayos de sol sobre las naves romanas e incendiarlas. Apolonio conocía también las propiedades de reflexión de las restantes secciones cónicas, por ejemplo, que todos los rayos que emanan de un foco de un espejo en forma de elipsoide se reflejan sobre el otro foco. Da las propiedades más importantes de la elipse y la hipérbola en el libro III de sus *Secciones Cónicas* (cap. 4, sec. 12). Griegos posteriores, en particular Pappus, conocían evidentemente la propiedad focal del paraboloide.

El fenómeno de la refracción de la luz, es decir, la curvatura de los rayos de luz cuando ésta pasa por un medio cuyas propiedades cambian constantemente, o el repentino cambio de dirección de un rayo de luz cuando pasa de un medio a otro, como por ejemplo del aire al agua, fue estudiado por los griegos alejandrinos. Ptolomeo observó el efecto de la refracción en la atmósfera sobre rayos proce-

dentes del Sol y las estrellas e intentó, infructuosamente, encontrar la ley correcta de la refracción cuando la luz pasa del aire al agua o del aire al cristal. Su *Optica*, que trata de los espejos y la refracción, se ha conservado.

8. La Astrología

Aunque actualmente la Astrología no es aceptada como ciencia, en civilizaciones antiguas tenía esta consideración. La astrología desarrollada por los griegos alejandrinos de alrededor del siglo II a. C. era diferente de la astrología babilonia del período asirio. La última se dedicaba exclusivamente a sacar conclusiones sobre el rey y asuntos de Estado a partir de observaciones de la posición de los planetas. No había ningún cálculo y la apariencia del cielo en el momento del nacimiento no jugaba ningún papel. Sin embargo, la astrología helénica o la alejandrina eran personales: predecían el futuro y el destino de personas concretas basándose en las posiciones calculadas del Sol, la Luna y los cinco Planetas del Zodiaco en el momento del nacimiento. Para evaluar estos datos se construyó un enorme cuerpo de doctrinas.

Ciertamente, esta ciencia fue tomada en serio por los griegos alejandrinos. Ptolomeo escribió un trabajo muy conocido sobre la cuestión, el *Quatripartite* o *Tetrabiblos*, o *Cuatro Libros Acerca de la Influencia de las Estrellas*, en el que da reglas para las predicciones astrológicas que fueron utilizadas durante un millar de años.

La importancia de la astrología en la historia de las ciencias radica en que motivó el estudio de la Astronomía, no sólo en Grecia sino también en la India, Arabia y en la Europa medieval. La Astrología fomentó la Astronomía en mucho mayor medida que la Alquimia lo hizo con la Química. Curiosamente, los errores en las predicciones astrológicas eran atribuidos a errores en Astronomía y no a lo incierto de las doctrinas astrológicas.

La Grecia alejandrina presentó los comienzos de la aplicación de la matemática a la medicina, de forma peculiar en gran medida a través de la astrología. Los doctores, llamados «iatromatemáticos», empleaban signos astrológicos para decidir los tratamientos. Galeno, el gran médico de los tiempos griegos, era un firme creyente en la astrología, quizá por ello es disculpable que Ptolomeo, el astrónomo más renom-

brado, también lo fuera. Esta conexión entre la matemática y la medicina se reforzó en la Edad Media.

Nuestra narración de la ciencia griega, en lo que concierne a la matemática, ha tratado sobre las ciencias matemáticas. Los griegos se dedicaron a otras investigaciones en áreas en las que la matemática, al menos en aquella época, no jugaba ningún papel. Además, mejoraron experimentos y llevaron a cabo observaciones, estas últimas especialmente en Astronomía. Sin embargo, su logro fundamental es que dieron un gran valor a las matemáticas dentro del contexto científico. El diálogo Platónico *Philebo* expresaba en primer lugar el pensamiento de que cada ciencia es una ciencia solamente en la medida en que contiene matemáticas; este principio ganó muchos adeptos a partir de los logros de los griegos. Además, los griegos pusieron claramente en evidencia que la Naturaleza está diseñada matemáticamente. Fue su visión de la Naturaleza y su iniciación de la investigación matemática de la Naturaleza lo que inspiró la creación de la matemática, en la época griega y en todos los siglos sucesivos.

Bibliografía

Apostle, H. G.: *Aristotle's Philosophy of Mathematics*, University of Chicago Press, 1952.
Berry, Arthur: *A Short History of Astronomy*, Dover (reimpresión), 1961, caps. 1-2.
Clagett, Marshall: *Greek Science in Antiquity*, Abelard-Schuman, 1955.
Dreyer, J. L. E.: *A History of Astronomy from Thales to Kepler*, Dover (reimpresión), 1953, caps. 1-9.
Farrington, Benjamin: *Ciencia griega*, Barcelona, Icaria, 1986.
Gomperz, Theodor: *Greek Thinkers*, 4 vols. John Murray, 1920.
Heath, Thomas L.: *Greek Astronomy*, J. M. Dent and Sons, 1932.
—: *Aristarchus of Samos*, Oxford University Press, 1913.
—: *The Works of Archimedes*, Dover (reimpresión), 1953, pp. 189-220 y 253-300.
Jaeger, Werner: *Paideia*, México, FCE, 1990.
Pannekoek, A.: *A History of Astronomy*, John Wiley and Sons, 1961.
Sambursky, S.: *El mundo físico de los griegos*, Madrid, Alianza, 1990.
Santillana, G. de: *The Origins of Scientific Thought from Anaximander to Proclus, 600 a. C. to 300 d. C.* University of Chicago Press, 1961.
Sarton, George: *A History of Science*, Harvard University Press, 1952 y 1959. Vols. 1 y 2.

Ver Eecke, Paul: *Euclide, L'Optique et la catoptrique,* Albert Blanchard, 1959.
Wedberg, Anders: *Plato's Philosophy of Mathematics,* Almqvist and Wiksell, 1955.

Capítulo 8

EL FINAL DEL MUNDO GRIEGO

Quien comprenda a Arquímedes y Apolonio admirará menos los logros de los hombres más ilustres de tiempos posteriores.

G. W. LEIBNIZ

1. Reseña de las realizaciones griegas

Aunque la civilización greco-alejandrina perduró hasta el año 640 d. C., en el que finalmente fue destruida por los mahometanos, es evidente que, a causa de su productividad decreciente, la civilización había entrado ya en declive durante los primeros siglos de la era cristiana. Antes de entrar a considerar las razones de este declive resumiremos las realizaciones y las imperfecciones de la matemática griega y tomaremos nota de los problemas que ha dejado para generaciones futuras. Los griegos alcanzaron grandes metas, y la continuación de la matemática, cuando fue retomada por los europeos tras pequeñas incursiones a cargo de hindúes y árabes, estuvo tan completamente determinada por el legado de los griegos que es importante tener claro dónde se sitúa su matemática.

Los griegos se caracterizaron por hacer matemática abstracta. Esta contribución principal es de una relevancia y un valor inconmensurables por el hecho de que un mismo triángulo abstracto o una misma ecuación algebraica se puede aplicar a cientos de situaciones físicas diferentes, que es donde se ha demostrado que radica el secreto de la potencia de la matemática.

Los griegos insistieron en las demostraciones deductivas. Este fue sin duda un avance extraordinario. De los cientos de civilizaciones que habían existido, algunas habían desarrollado algún tipo rudimentario de aritmética y de geometría. Sin embargo, ninguna civilización, aparte de los griegos concibió la idea de establecer conclusiones exclusivamente a través del razonamiento deductivo. La decisión de exigir demostraciones deductivas está en contraposición absoluta con los métodos utilizados por el hombre hasta entonces en los demás campos; es, de hecho, casi irracional, porque casi todo el conocimiento altamente fiable se adquiría a través de la experiencia, la inducción, el razonamiento por analogía y la experimentación. Pero los griegos buscaban verdades y vieron que solamente las obtendrían por los métodos infalibles del razonamiento deductivo. Comprendieron también que para llegar a verdades seguras debían partir de verdades y estar seguros de no suponer ningún hecho no garantizado. Por tanto, establecieron todos sus axiomas de forma explícita y además adoptaron la práctica de situarlos muy al principio de sus trabajos para que de esta manera pudieran ser examinados de golpe con sentido crítico.

Después de concebir este plan enormemente importante para asegurar un conocimiento seguro, los griegos introdujeron una sofisticación que difícilmente podía esperarse de los innovadores. Su conciencia de que los conceptos no podían ser contradictorios y de que no se puede construir una estructura consistente trabajando con figuras no existentes (tales como un poliedro regular de diez caras) pone de manifiesto una agudeza de pensamiento casi sobrehumana y ciertamente sin precedentes. Como sabemos ahora, su método para establecer la existencia de los conceptos, con lo que podían trabajar con ellos, era demostrar que podían construirse con el uso de regla y compás.

La potencia de los griegos para intuir teoremas y demostraciones queda atestiguada por el hecho de que los *Elementos* de Euclides contienen 467 proposiciones y las *Secciones Cónicas* de Apolonio, 487, obtenidas todas ellas a partir de 10 axiomas enunciados en los *Elementos.* La coherencia que proporcionan las estructuras deductivas no ofrece ninguna duda, ni siquiera secundaria en cuanto a su importancia, ni quizá tampoco secundaria en cuanto a atención. Aún la posibilidad de obtener los mismos resultados a partir de numerosos conjuntos de axiomas distintos —si bien igualmente fiables— podría dar una versión del conocimiento menos manejable y asimilable.

La contribución griega al contenido de la matemática —geometría

plana y del espacio, trigonometría plana y esférica, los comienzos de la teoría de números, la ampliación de la aritmética y el álgebra de Egipto y Babilonia— es enorme, especialmente si se tiene en cuenta el reducido número de personas dedicadas a ellas y los escasos siglos a los que se extendió su actividad. A estas contribuciones debemos añadir el álgebra geométrica, que esperaba solamente el reconocimiento de los números irracionales y la instauración del lenguaje simbólico para convertirse en la base de gran parte del álgebra elemental. Por otra parte, el estudio de figuras curvilíneas por el método exhaustivo, a pesar de que formaba parte de su geometría, merece una mención especial ya que constituye el comienzo del cálculo.

Una contribución igualmente importante y un motivo de inspiración para generaciones posteriores fue la concepción griega de la naturaleza. Los griegos identificaban la matemática con la realidad del mundo físico y veían en ella la verdad última sobre la estructura y el plan del Universo. Encontraron la alianza entre la matemática y el estudio desinteresado de la naturaleza, lo que se ha convertido desde entonces en la gran base de la ciencia moderna. Además, fueron bastante más lejos a la hora de racionalizar la Naturaleza, al establecer la firme convicción de que el Universo está en efecto trazado matemáticamente, es controlable, está regido por leyes y es comprensible para el hombre.

El atractivo estético de la matemática no fue pasado por alto. En la época griega, las matemáticas eran consideradas también como un arte; la belleza, la armonía, la sencillez, la claridad y el orden eran reconocidas en ellas. La aritmética, la geometría y la astronomía se tomaban como el arte de la mente y la música, el del espíritu. Platón se complacía con la geometría; Aristóteles no separaría las matemáticas de la estética, pues el orden y la simetría eran para él elementos importantes de belleza, y éstos los encontraba en las matemáticas. Evidentemente, los intereses racionales y estéticos, así como los morales, son difícilmente separables en el pensamiento griego. Leemos una y otra vez que la esfera es el cuerpo con la forma más bella y es, por tanto, divina y buena. El círculo participaba junto con la esfera de esta llamada a la estética; parecía obvio por tanto que el círculo fuera el camino de aquellos cuerpos que representaban lo inmutable, el orden eterno del cielo, mientras que el movimiento lineal prevalecía sobre la tierra imperfecta. No hay duda que fue la estética de las matemáticas lo que dio lugar a que los matemáticos griegos prosiguie-

ran la exploración de temas concretos una vez utilizados para la comprensión del mundo físico.

2. Las limitaciones de la matemática griega

Pese a sus logros maravillosos, las matemáticas griegas eran defectuosas. Sus limitaciones señalan los caminos del progreso al que, sin embargo, todavía no estaban abiertas.

La primera limitación fue la incapacidad para admitir el concepto de número irracional. Esto significaba no solamente una restricción de la aritmética y el álgebra, sino también una vuelta a la geometría y el énfasis en ella, ya que el pensamiento geométrico evitaba una presentación explícita de lo irracional como un número. Si los griegos hubieran afrontado el número irracional podrían haber adelantado el desarrollo de la aritmética y el álgebra; e incluso, si ellos mismos no lo hubieran hecho, no habrían impedido que lo hicieran generaciones posteriores, que fueron inducidas a pensar que solamente la geometría ofrecía un fundamento seguro para el estudio de magnitudes cuyos valores podían incluir irracionales. Arquímedes, Herón y Ptolomeo comenzaron a trabajar con los irracionales como números, pero no modificaron el carácter de las matemáticas griegas ni la impronta subsiguiente del pensamiento griego. El hecho de que los griegos se concentraran en la geometría nubló la visión de otras generaciones al enmascarar la correspondencia íntima entre los conceptos geométricos y los aritméticos y las operaciones. El fracaso a la hora de definir, aceptar y conceptualizar los irracionales como números forzó una distinción entre número y magnitud. En consecuencia, el álgebra y la geometría fueron contempladas como disciplinas sin ninguna relación mutua.

De haber estado menos dedicados a ser lógicos y rigurosos, los griegos podían haber aceptado (y operado con) los números irracionales, igual que lo hicieron los babilonios y otras civilizaciones que sucedieron a los griegos. Pero la base intuitiva de la idealización no estaba clara, y la construcción lógica no entraba de lleno en sus poderes. La virtud de los griegos de insistir en la exactitud de los conceptos y las definiciones constituía un defecto en lo que concierne a las matemáticas creativas.

La restricción del rigor matemático a la geometría (además de a la teoría de números) dio lugar a otra desventaja importante: el uso de

métodos geométricos condujo a demostraciones cada vez más complicadas a medida que las matemáticas se iban ampliando, particularmente en el área de la geometría del espacio. Además, incluso en las demostraciones más sencillas, hay una ausencia de métodos generales, lo cual es claro para nosotros ahora por estar en posesión de la geometría analítica y del cálculo. Cuando se consideran las dificultades que encontró Arquímedes para hallar el área de un segmento parabólico o el área subtendida por un arco de su espiral, y se compara esto con los métodos modernos de cálculo, se aprecia la efectividad de estos últimos.

Los griegos no sólo restringieron las matemáticas en gran medida a la geometría, sino que incluso limitaron esta disciplina a las figuras que se podían obtener a partir de la línea recta y el círculo. De acuerdo con esto, las únicas superficies admitidas eran aquellas que se podían obtener haciendo girar líneas rectas y círculos alrededor de un eje, como por ejemplo el cilindro, el cono y la esfera, formados por la revolución de un rectángulo, un triángulo y un círculo, respectivamente, alrededor de una recta; el prisma, que es un cilindro especial, y la pirámide, que resulta de la descomposición de un prisma. Las secciones cónicas se introdujeron al cortar conos mediante un plano. Curvas como la cuadratriz de Hipias, la concoide de Nicomedes y la cisoide de Diocles quedaron como algo marginal de la geometría; recibieron, en este caso, el calificativo de mecánicas, más que geométricas.

La clasificación de las curvas a cargo de Pappus es un intento de mantener unos límites fijos. Los griegos, conforme a los criterios de Pappus, distinguían las curvas como sigue: los lugares planos o curvas planas eran los que se podían construir a partir de líneas rectas y círculos; las cónicas recibían el nombre de lugares sólidos puesto que se originaban a partir del cono; las curvas lineales, como cuadratrices, concoides, cisoides y espirales formaban la tercera clase. Análogamente, distinguían entre problemas planos, sólidos y lineales. Los problemas planos se resolvían mediante rectas y círculos; los problemas sólidos, a través de una o más secciones cónicas. Los problemas que no podían resolverse por medio de líneas rectas, círculos o cónicas se llamaban lineales, debido a que utilizaban líneas (curvas) que tenían un origen más complicado o menos natural que las anteriores. Pappus destacó la importancia de resolver problemas mediante lugares planos o sólidos ya que entonces se podía dar el criterio para una solución efectiva.

¿Por qué los griegos limitaron su geometría a la recta, el círculo y a figuras directamente derivadas de ellos? Una razón es que de esta manera resolvían el problema de determinar la existencia de figuras geométricas. Como ya hemos visto, Aristóteles, en particular, señalaba que debemos estar seguros de que los conceptos introducidos no son autocontradictorios; es decir, hemos de demostrar que existen. Para aclarar este punto, los griegos, al menos en un principio, admitían exclusivamente aquellos conceptos que se podían construir. La recta y el círculo se admitían como constructibles en los postulados, pero las demás figuras debían poderse construir con la recta y el círculo.

Sin embargo, el uso de construcciones para determinar su existencia no se aplicaba a figuras tridimensionales. En este punto, los griegos aceptaban, en apariencia, lo que era intuitivamente claro, como por ejemplo la existencia de figuras de revolución tales como la esfera, el cilindro y el cono. Las secciones planas de estas figuras daban lugar a curvas tales como las secciones cónicas; así, fueron aceptadas incluso figuras planas cuya existencia no había sido establecida —aunque con reticencias—. Descartes hace notar esta cuestión muy al principio del libro II de *La Géometrie:* «Es cierto que las secciones cónicas no fueron aceptadas nunca de buena gana en la geometría antigua...»

Otro motivo para la restricción a la recta, el círculo y otras figuras derivadas de ellos parte de Platón, ya que de acuerdo con sus ideas tenía que estar claro lo que era aceptable. Mientras el número entero parecía ser aceptable como una idea clara en sí misma, pese a que nunca fue explícitamente definida por los griegos, las figuras geométricas tenían que construirse con precisión. Rectas y círculos, así como figuras que se derivan de ellos estaban claros, mientras que las curvas introducidas mediante instrumentos mecánicos (distintos de la regla y el compás) no lo estaban, por lo que eran inadmisibles. La restricción a figuras claramente definidas dio lugar a una geometría simple, ordenada, armoniosa y bella.

Al insistir en una unidad, una completitud y una sencillez para su geometría y al separar el pensamiento especulativo de la utilidad, la geometría clásica griega limitó sus logros. Restringió la visión de las gentes y cerró sus mentes a nuevos pensamientos y métodos. Llevaba en sí misma la semilla de su propia muerte. La estrechez de su campo de acción, la exclusividad de su punto de vista y la demanda estética sobre ella pudieron haber detenido su evolución, si no fuera por las

influencias de la civilización alejandrina, que ensanchó las perspectivas de los matemáticos griegos.

Las doctrinas filosóficas griegas limitaron las matemáticas en otra dirección. A lo largo de todo el período clásico, creían que el hombre no creaba los hechos matemáticos: preexistían. El hombre se limitaba a descubrirlos y a registrarlos. En el *Teeteto* Platón compara la búsqueda del conocimiento a un pájaro cautivo encerrado en una jaula. Los pájaros, ya prisioneros, necesitan solamente ser cogidos con la mano. Esta creencia acerca de la naturaleza de las matemáticas no prevaleció.

Los griegos no consiguieron comprender lo infinitamente grande, lo infinitamente pequeño y los procesos infinitos. Ellos «se atemorizaban ante el silencio de los espacios infinitos». Los pitagóricos asociaron lo bueno y lo malo con lo limitado y lo ilimitado respectivamente. Aristóteles dice que el infinito es imperfecto, inacabado y en consecuencia inabordable; no tiene forma y es confuso. Los objetos tienen una naturaleza únicamente cuando están delimitados y son distinguibles.

Para evitar cualquier afirmación acerca de la infinitud de la línea recta, Euclides dice que un segmento lineal (utiliza la palabra «línea» en este sentido) puede prolongarse todo lo que sea necesario. La reticencia para incluir lo infinitamente grande puede verse también en el enunciado del axioma de las paralelas de Euclides. En vez de considerar dos líneas que se prolongan indefinidamente y dar una condición directa o una hipótesis bajo la cual las rectas paralelas podrían existir, su axioma de las paralelas da una condición por la cual dos líneas rectas se cortan en algún punto finito.

El concepto de lo infinitamente pequeño está implícito en la relación existente entre los puntos de una línea o la relación entre lo discreto y lo continuo; las paradojas de Zenón pudieron haber sido la causa de que los griegos dejaran de lado esta cuestión. La relación entre punto y recta abrumaba a los griegos y llevó a Aristóteles a separar ambos conceptos. Pese a que admite que los puntos estaban sobre rectas, dice que una recta no puede estar formada de puntos y que lo continuo no puede construirse a partir de lo discreto (cap. 3, sec. 10). Esta distinción contribuyó también a la presunta necesidad de separar el número de la geometría, ya que los números eran discretos, mientras que la geometría trataba de magnitudes continuas.

Puesto que recelaban de los procesos infinitos, omitieron el proceso de paso al límite. Al aproximar un círculo mediante un

polígono se contentaban con hacer que la diferencia fuera menor que cualquier cantidad dada previamente, pero se exigía que fuera siempre estrictamente positiva. De esta manera el proceso queda claro para la intuición; el paso al límite, por otra parte, habría llevado consigo la consideración de lo infinitamente pequeño.

3. Los problemas legados por los griegos

Las limitaciones del pensamiento matemático griego conducen de manera casi automática a los problemas que dejaron para las generaciones futuras. El fracaso a la hora de aceptar los irracionales como números dejó ciertamente abierta la cuestión de si se podía asignar un número a razones inconmensurables, con lo que éstas podrían estudiarse desde el punto de vista de la aritmética. Con el número irracional, el álgebra se ampliaría también. En vez de regresar a la geometría para resolver ecuaciones cuadráticas, o de otro tipo, que podían tener raíces irracionales, estos problemas se podrían abordar en términos numéricos y el álgebra se desarrollaría a partir de la situación en que la dejaron los egipcios y los babilonios o donde la dejó Diofanto, que rechazó la idea de considerar los irracionales como números.

Incluso para los números enteros y las razones de números enteros, los griegos no tenían ninguna base lógica; la sustituyeron por algunas definiciones bastante vagas, establecidas por Euclides en los libros VII y IX de los *Elementos*. La necesidad de un fundamento lógico del sistema de números se vio acrecentada por el uso libre de los números, incluidos los irracionales, por parte de los alejandrinos; a este respecto continuaron estrictamente las tradiciones empíricas de egipcios y babilonios. Por tanto, los griegos legaron dos ramas de las matemáticas completamente distintas y desigualmente desarrolladas. Por una parte estaba la rigurosa, deductiva y sistemática geometría y por otra, la heurística y empírica aritmética y su extensión al álgebra.

La incapacidad para la construcción de un álgebra deductiva significa que el rigor matemático quedó confinado a la geometría; de hecho, éste siguió siendo el caso hasta los siglos XVII y XVIII, cuando el álgebra y el cálculo ya se habían extendido. Incluso entonces se entendía todavía que las matemáticas rigurosas se referían a la geometría.

La restricción de la geometría euclídea a conceptos que se pudie-

ran construir con regla y compás dejó dos tareas a las matemáticas. La primera era específica: probar la cuadratura del círculo, la trisección del ángulo y la duplicación del cubo con la regla y el compás. Estos tres problemas ejercieron una gran fascinación e incluso hoy en día llaman la atención a la gente, pese a que, como veremos, estaban resueltos en el siglo XIX.

La segunda tarea era ampliar los criterios para la existencia. La posibilidad de ser construido como medio de probar la existencia se convirtió en algo excesivamente restrictivo para los conceptos con los que iban a trabajar las matemáticas (y con los que más tarde lo hicieron). Además, como algunas longitudes no se pueden construir, la recta euclídea es incompleta; es decir, no contiene, en sentido estricto, las longitudes no constructibles. Para ser internamente completas y más útiles al estudio del mundo físico, las matemáticas debían liberarse a sí mismas de una limitación técnica para el establecimiento de la existencia de los conceptos.

Como vimos, el intento de evitar una afirmación directa acerca de líneas rectas paralelas infinitas hizo que Euclides enunciara el axioma de la paralelas de una forma mucho más complicada. Consiguió que, al hablar de esta manera, este axioma perdiera la autoevidencia de los nueve restantes y hay buenas razones para pensar que evitó usarlo mientras pudo. Varios griegos intentaron encontrar axiomas que sustituyeran al de las paralelas, o probarlo en función de los otros nueve. Ptolomeo escribió acerca de esta cuestión; Proclo, en su comentario sobre Euclides, da el intento de Ptolomeo de demostrar el postulado de las paralelas e intenta a su vez probarlo por sí mismo. Simplicio cita otros dos investigadores y añade que la gente «en la antigüedad» puso objeciones al uso del axioma de las paralelas.

Estrechamente relacionada con el problema del postulado de las paralelas está la cuestión de saber si el espacio físico es infinito. Euclides supone en el postulado 2 que un segmento de línea recta puede extenderse tanto como sea preciso; usa este hecho, pero solamente para obtener grandes longitudes finitas —por ejemplo en el libro I, proposiciones 11, 16 y 20—. Herón da nuevas demostraciones de estos teoremas y evita prolongar las líneas, con el fin de salir al paso de las objeciones de cuantos negaran que el espacio se podía abarcar por extensión. Aristóteles había considerado la cuestión de averiguar si el espacio era infinito y dio seis argumentos de naturaleza no matemática para probar que es finito; pronosticaba que esta cuestión sería problemática.

Otro problema importante legado a la posteridad fue el cálculo de áreas limitadas por curvas y volúmenes limitados por superficies. Los griegos, especialmente Eudoxo y Arquímedes, no solamente habían abordado la cuestión sino que, como hemos visto, lograron progresos considerables usando el método de exhausción. Pero el procedimiento presentaba dificultades como mínimo en dos aspectos: en primer lugar, cada problema requería algún esquema ingenioso para aproximar el área o el volumen en cuestión; sin embargo, la inventiva humana simplemente no disponía de suficientes recursos para las áreas y volúmenes que tenía que calcular después. En segundo lugar, el resultado al que llegaban los griegos consistía habitualmente en probar la equivalencia del área o volumen deseados con el área o el volumen de alguna figura más sencilla cuya medida todavía no era conocida cuantitativamente. Pero es precisamente este conocimiento cuantitativo el que requieren las aplicaciones.

4. La desaparición de la civilización griega

Comenzando aproximadamente con el principio de la era cristiana, la vitalidad de la actividad matemática griega declinó rápidamente. Las únicas contribuciones importantes de la nueva era fueron las de Ptolomeo y Diofanto. Los grandes comentaristas Pappus y Proclo merecen también la atención, pero en realidad son los que cierran la nómina. El declive de esta civilización, que durante cinco o seis siglos aportó contribuciones que sobrepasaban en gran medida, tanto en extensión como en brillantez, las de cualquier otra, requiere una explicación.

Desgraciadamente, los matemáticos están sujetos a los designios de la historia, igual que el último labrador. Basta con familiarizarse con los hechos más superficiales de la historia política de Alejandría para darse cuenta de que no sólo las matemáticas, sino cualquier tipo de actividad cultural, estaban destinadas a sufrir. Mientras la civilización greco-alejandrina estuvo gobernada por los Ptolomeos, floreció. El primer desastre fue el advenimiento de los romanos, cuyo único papel en la historia de las matemáticas fue el de agentes de destrucción.

Antes de discutir su impacto sobre la civilización greco-alejandrina, veamos algunos hechos acerca de las matemáticas en Roma y la naturaleza de la civilización romana: las matemáticas romanas apenas si son dignas de mención. El período durante el cual

los romanos figuran en la historia comprende los años que van desde aproximadamente el 750 a. C. hasta el 476 de nuestra era, más o menos el mismo tiempo durante el cual floreció la civilización griega. Además, como veremos, a partir del 200 a. C. los romanos estuvieron en estrecho contacto con los griegos. Con todo, en los once siglos no hubo ningún matemático romano; además de otros detalles este hecho habla virtualmente por sí mismo de toda la historia de las matemáticas en Roma.

Los romanos tenían una aritmética rudimentaria y algunas fórmulas geométricas aproximadas que posteriormente fueron complementadas por copias de las greco-alejandrinas. Sus símbolos para los números enteros nos son familiares. Para calcular con números enteros utilizaban diversos tipos de ábacos. Los cálculos se hacían también con los dedos y con la ayuda de tablas especialmente preparadas.

Las fracciones en Roma estaban en base 12. Se usaban símbolos y palabras especiales para designar 1/12, 2/12, ..., 11/12, 1/24, 1/36, 1/48, 1/96, ... El origen de la base 12 puede ser la relación existente entre el mes lunar y el año. La unidad de peso, por cierto, era el *as*; un doceavo del mismo era la *uncia,* de la que derivan nuestras onza y pulgada.

El principal uso de la aritmética y la geometría en Roma fue la agrimensura, para determinar las fronteras de las ciudades y para medir terrenos para las casas y los templos. Los agrimensores calculaban la mayoría de las cantidades que precisaban usando solamente instrumentos sencillos y triángulos congruentes.

Debemos a los romanos una mejora del calendario. En los tiempos de Julio César (100-44 a. C.) el año básico romano tenía 12 meses, que totalizaban 355 días. En años alternos se añadía un mes intercalado de 22 ó 23 días de manera que el año promedio tenía 366 días y 1/4. Para mejorar este calendario, César llamó a Sosígenes, un alejandrino, que aconsejó un año de 365 días con un año bisiniesto cada cuatro años. El calendario Juliano fue adoptado el año 45 a. C.

A partir del año 50 a. C., aproximadamente, los romanos escribieron sus propios libros técnicos; todo el material de base, sin embargo, se tomó de las fuentes griegas. El más famoso de estos trabajos técnicos son los diez libros de Vitrubio sobre arquitectura, que datan del año 14 a. C. Aquí, también, el material es griego. Es curiosa la afirmación de Vitrubio de que los tres grandes descubrimientos matemáticos son el triángulo rectángulo 3, 4, 5, la irracionalidad de la

diagonal del cuadrado unidad y la solución de Arquímedes del problema de la corona. Da otros hechos que implican el uso de matemáticas, tales como las proporciones de las partes del cuerpo humano ideal, algunas relaciones aritméticas armónicas y relaciones aritméticas acerca de las capacidades de las catapultas.

Entre los romanos el término «matemáticas» cayó en desgracia a causa de que los astrólogos recibían el nombre de *mathematicii,* y la astrología fue condenada por los emperadores romanos. El emperador Diocleciano (245-316 de nuestra era) hacía distinciones entre geometría y matemáticas. La primera se enseñaba y aplicaba en las escuelas públicas; pero el «arte de las matemáticas» —esto es, la astrología— fue condenado y prohibido completamente. El «código de matemáticas y malas artes», la ley romana que prohibía la astrología, se aplicó también en Europa durante la Edad Media. Sin embargo, los emperadores romanos y los cristianos empleaban astrólogos en sus cortes por la posibilidad de que pudiera haber algo de cierto en sus profecías. La distinción entre los términos «matemático» y «geómetra» duró hasta bien pasado el Renacimiento. Incluso en los siglos XVII y XVIII, «geómetra» significaba lo que hoy entendemos por «matemático».

Los romanos eran un pueblo práctico y hacían alarde de su practicismo. Diseñaron y completaron grandes proyectos de ingeniería —viaductos, magníficas vías que sobreviven todavía hoy, puentes, edificios públicos y mediciones de terrenos— pero se negaron a considerar cualquier idea que pudiera haber detrás de las aplicaciones particulares y concretas que estaban realizando en aquel momento. La actividad romana acerca de las matemáticas viene dada por Cicerón: «Los griegos dieron al geómetra el más alto honor; de acuerdo con esto, nada tenía un progreso más brillante que las matemáticas. Pero nosotros hemos establecido como límite de este arte su utilidad para medir y contar.»

Los emperadores romanos no dieron apoyo a las matemáticas tal como hicieron los Ptolomeos en Egipto. Ni los romanos comprendían la ciencia pura. Su incapacidad para desarrollar las matemáticas es notoria, debido a que gobernaban un ancho imperio y porque lo que buscaban era la resolución de problemas prácticos. La lección que se puede aprender de la historia de los romanos es que los pueblos que desdeñan los trabajos de matemáticos y científicos altamente teóricos y desacreditan su utilidad ignoran la forma en la que se han presentado importantes desarrollos prácticos.

Volvamos de nuevo al papel que jugaron los romanos en la historia política y militar de Grecia. Tras haber asegurado el control del centro y el norte de Italia, conquistaron las ciudades griegas del sur de Italia y Sicilia. (Recordemos que Arquímedes contribuyó a la defensa de Siracusa cuando los romanos atacaron la ciudad y murió a manos de un soldado romano.) Los romanos conquistaron Grecia propiamente dicha el año 146 a. C. y Mesopotamia el 64 a. C. Al intervenir en las luchas internas de Egipto entre Cleopatra, la última de la dinastía Ptolomea, y su hermano, César manipuló para asegurarse un dominio sobre el país. El año 47 a. C., César prendió fuego a la flota egipcia que navegaba y estaba anclada en el puerto de Alejandría; el fuego se extendió a la ciudad e incendió la Biblioteca. Dos siglos y medio de recolección de libros y medio millón de manuscritos, que representaban el esplendor de la antigua cultura, fueron borrados. Afortunadamente un excedente de libros que no habían podido ser colocados en la repleta Biblioteca estaban en aquellos tiempos almacenados en el templo de Serapis y éstos no fueron incendiados. Asimismo, Atalo III de Pérgamo, que murió el 133 a. C., había legado a Roma su gran colección de libros. Marco Antonio regaló esta colección a Cleopatra y se sumaron a los libros del templo. La colección resultante volvió a ser enorme de nuevo.

Los romanos regresaron a la muerte de Cleopatra, el año 31 a. C., y a partir de este momento controlaron Egipto. Su interés en extender su poder político no incluía la difusión de su cultura. Las áreas subyugadas se convirtieron en colonias, de las que se extraía una gran riqueza mediante la expropiación y los impuestos. Como la mayoría de los emperadores romanos eran propietarios, arruinaban todos los países que controlaban. Cuando se producía algún levantamiento, como ocurrió, por ejemplo, en Alejandría, los romanos no dudaban en matar de hambre a la población y, una vez dominada la revuelta, matar a miles de habitantes.

La historia del final del imperio romano es también relevante. El emperador Teodosio (gobernó entre el 379 y el 395) dividió su ancho imperio entre sus dos hijos, Honorio, que fue el que gobernó Italia y Europa occidental, y Arcadio, que gobernó Grecia, Egipto y el Oriente próximo. La parte occidental fue conquistada por los godos durante el siglo V y su historia posterior pertenece ya a la de la Europa medieval. La parte oriental, que incluía Egipto (durante un tiempo), Grecia y lo que en la actualidad es Turquía, conservó su independencia hasta que fue conquistada por los turcos el año 1453. Puesto que el

Imperio Romano de Oriente, conocido también como el Imperio Bizantino, incluía Grecia propiamente dicha, la cultura y las obras griegos fueron conservados en alguna medida.

Desde el punto de vista de la historia de las matemáticas, la aparición del cristianismo tuvo consecuencias poco afortunadas. Pese a que los jefes cristianos adoptaron varios mitos y costumbres griegas y orientales con la intención de hacer el cristianismo más aceptable a los conversos, se opusieron a las enseñanzas paganas y ridiculizaron las matemáticas, la astronomía y la física; se prohibió a los cristianos contaminarse con las enseñanzas griegas. A pesar de la persecución cruel de que fueron objeto por parte de los romanos, el cristianismo se difundió y llegó a tener tal importancia que el emperador Constantino (272-337) se vio obligado a adoptarlo como la religión oficial del Imperio Romano. Así, los cristianos fueron capaces de llevar a cabo una mayor destrucción de la cultura griega. El emperador Teodosio proscribió las religiones paganas y, en 392, dio la orden de que los templos griegos fueran destruidos. Muchos de ellos fueron convertidos en iglesias, a pesar de que a menudo estaban adornados todavía con esculturas griegas. Los paganos fueron atacados y asesinados por todo el Imperio. El destino de Hipatia, una matemática alejandrina de relevancia, e hija de Teón de Alejandría, simboliza el fin de la era. Como consecuencia de haberse negado a abandonar la religión griega, los fanáticos cristianos la apresaron en las calles de Alejandría y la despedazaron.

Los libros griegos fueron quemados a millares. El año en que Teodosio prohibió las religiones paganas, los cristianos destruyeron el templo de Serapis, que todavía albergaba la única gran colección de obras griegas. Se estima que fueron destruidos 300.000 manuscritos. Muchos más trabajos escritos en pergamino fueron requisados por los cristianos y usados para sus propios escritos. El año 529 el emperador romano de Oriente, Justiniano, cerró todas las escuelas griegas de filosofía, incluida la Academia de Platón. Muchos sabios griegos abandonaron el país y algunos —por ejemplo, Simplicio— se asentaron en Persia.

El último suspiro para Alejandría fue la conquista de Egipto por los rebeldes mahometanos el año 640. Los libros que todavía quedaban fueron destruidos basándose en la proclama dada por Omar, el conquistador árabe: «Los libros, o bien contienen lo que ya está en el Corán, en cuyo caso no tenemos que leerlos, o bien contienen lo contrario de lo que está en el Corán, en cuyo caso no debemos

leerlos.» Como consecuencia de esto los baños de Alejandría se calentaron con el fuego de los rollos de pergamino.

Tras la captura de Alejandría por los mahometanos, la mayoría de los sabios emigraron a Constantinopla, que se había convertido en la capital del Imperio Romano de Oriente. Pese a que no florecería ninguna actividad respecto de las líneas del pensamiento griego en la atmósfera cristiana hostil de Bizancio, este flujo de eruditos y sus trabajos de relativa calidad incrementaron el tesoro del conocimiento que llegó hasta Europa ochocientos años después.

Resulta quizá fuera de lugar contemplar lo que podía haber sido. Pero no se puede dejar de considerar que la civilización greco-alejandrina terminó su activa vida científica en los umbrales de la era moderna. Fue la poco habitual combinación de intereses teóricos y prácticos que demostró ser fecunda mil años después. Durante los últimos siglos de su existencia, gozaron de libertad de pensamiento, que es también esencial para una cultura floreciente y abordaron y llevaron a cabo avances de gran relevancia en varios campos que fueron los que se convirtieron en primordiales durante el Renacimiento: geometría cuantitativa plana y del espacio, trigonometría, álgebra, cálculo y astronomía.

Se dice a menudo que el hombre propone y Dios dispone. Es más acertado decir que los griegos y Dios propusieron y el hombre dispuso. Los matemáticos griegos fueron borrados, pero el fruto de su trabajo llegó hasta Europa por el camino que vamos a relatar.

Bibliografía

Cajori, Florian: *A History of Mathematics*, Macmillan, 1919, pp. 63-68.

Gibbon, Edward: *Historia de la decadencia y ruina del Imperio Romano*, Madrid, Turner, 1984.

Parsons, Edward Alexander: *The Alexandrian Library*, The Elsevier Press, 1952.

Capítulo 9

LA MATEMATICA DE LOS HINDUES Y DE LOS ARABES

> Así como el Sol eclipsa las estrellas por su brillantez, también el hombre culto eclipsará la fama de otros en asambleas del pueblo si propone problemas algebraicos y todavía más si los resuelve.
>
> BRAHMAGUPTA

1. La primera matemática hindú

Los sucesores de los griegos en la historia de la matemática fueron los hindúes de la India. Pese a que las matemáticas indias llegaron a tener relevancia sólo después de recibir la influencia de los resultados griegos, en un principio hubo desarrollos autóctonos sin ninguna importancia.

La civilización hindú data como mínimo del 2000 a. C., pero, dentro de lo que podemos saber, no existía ningún tipo de matemáticas antes del 800 a. C. Durante el período de Sulvasutra, que va del 800 a. C. al 200 de nuestra era, los indios produjeron alguna matemática rudimentaria. No existía ningún documento matemático aislado pero se puede considerar algunos hechos recogidos de otros escritos, y de monedas e inscripciones.

A partir del siglo III a. C. aproximadamente, aparecen símbolos numéricos, que variaban considerablemente de un siglo a otro. Son típicos los símbolos de Brahmi:

—	=	≡							α	o				
1	2	3	4	5	6	7	8	9	10	20	30	40	50	60.

Lo que llama la atención en este conjunto es la existencia de un símbolo individual para cada uno de los números comprendidos entre 1 y 9. Sin embargo, no existía ningún símbolo para el cero y ningún tipo de notación posicional. El acierto de usar símbolos independientes no fue previsto indudablemente por el pueblo matemáticamente ignorante; la práctica pudo haber surgido a partir de la utilización de las primeras letras de las palabras que designaban dichos números.

Entre los escritos religiosos había una clase llamada Sulvasutras (reglas de la cuerda) que contenía instrucciones para la construcción de altares. En uno de los Sulvasutras del siglo IV o V a. C. se da una aproximación de $\sqrt{2}$, pero no hay ninguna indicación de que es precisamente una aproximación. Asimismo, no se conoce casi nada acerca de la aritmética de ese período.

La geometría de esta antigua época hindú se conoce algo mejor. Las reglas contenidas en los Salvasutras dan condiciones para las formas y tamaños de los altares. Las tres formas más comúnmente usadas eran el cuadrado, el círculo y el semicírculo; y fuera la que fuese la forma utilizada, el área tenía que ser la misma. Por tanto, los hindúes habían construido círculos que tenían la misma área que los cuadrados, o dos veces mayor para que se pudiera utilizar el semicírculo. Otra forma empleada era el trapecio isósceles; aquí se permitió usar una forma semejante y en consecuencia, aparecieron problemas geométricos adicionales para construir la figura semejante.

Al diseñar los altares autorizados los hindúes adquirieron algún conocimiento de los hechos geométricos básicos, como el teorema de Pitágoras, dado en la forma: «La diagonal de un cuadrilongo (rectángulo) produce por sí mismo las dos áreas a que dan lugar por separado cada uno de los lados del cuadrilongo. En general, la geometría de este período consiste en un conjunto inconexo de reglas verbales aproximadas para el cálculo de áreas y volúmenes. Apastamba (siglos IV o V a. C.) da una construcción para la obtención de un círculo con la misma área que un cuadrado, la cual usa, efectivamente, el valor 3,09 para π, pero él pensaba que la construcción era exacta. En toda la geometría de este período primitivo no se encuentra ninguna demostración; las reglas eran empíricas.

2. Aritmética y álgebra indias del período 200-1200

La segunda época de las matemáticas indias, el período alto, se puede datar groseramente desde el año 200 de nuestra era hasta el 1200. Durante la primera parte del mismo, la civilización de Alejandría influyó decisivamente en los indios. Varahamihira (c. 500), un astrónomo, dice: «Los griegos, pese a ser impuros [cualquiera que tenga una creencia diferente es impuro], deben ser honrados, puesto que fueron adiestrados en las ciencias y allí sobresalieron por encima de los demás. ¿Qué se puede decir, pues, de un brahmán si él une a su pureza la altura de la ciencia?» La geometría de los indios era realmente griega, pero ellos tenían un don especial para la aritmética. Igual que con el álgebra, pudieron haber partido de Alejandría y posiblemente de Babilonia, pero aquí, además, llegaron muy lejos en su desarrollo. La India estaba también un poco en deuda con China.

Los matemáticos más importantes del segundo período son Aryabhata (nacido el 476), Brahmagupta (nacido el 598), Mahavira (siglo IX) y Bhaskara (nacido el 1114). Muchos de sus trabajos y en general los de los matemáticos indios estaban motivados por la astronomía y la astrología. En realidad, no hay textos de matemáticas independientes, el material matemático aparece en capítulos de libros de astronomía.

Los métodos indios para la escritura de números en el año 600 eran numerosos y en algunos casos incluían palabras o sílabas para los símbolos numéricos. El año 600 se volvió otra vez a los antiguos símbolos brahmi, a pesar de que la forma concreta de los mismos varió a lo largo del período. La notación posicional en base 10, que había sido de uso limitado durante unos cien años, se convierte ahora en habitual. También el cero, que los griegos alejandrinos usaban en los primeros tiempos solamente para designar la ausencia de un número, fue considerado un número a todos los efectos. Mahavira dice que la multiplicación de un número por cero da cero y que la sustracción de cero no disminuye el valor del número. Sin embargo, afirma también que si se divide un número por cero, su valor permanece invariable. Bhaskara, al hablar de una fracción cuyo denominador es cero dice que dicha fracción permanece invariable aunque se añada o sustraiga cualquier cantidad, así como no sufre ningún cambio la inmutable divinidad cuando se crean y destruyen los mundos. Un número dividido por cero, añade, se designa como una cantidad infinita.

Para las fracciones en astronomía los indios usaban la notación posicional sexagesimal. Para otras finalidades empleaban una razón de enteros, pero sin la barra, como por ejemplo: $\begin{matrix}3\\4\end{matrix}$.

Las operaciones aritméticas eran muy parecidas a las nuestras. Por ejemplo, Mahavira da nuestra regla de división por una fracción: invierte y multiplica.

Los indios introdujeron los números negativos para indicar deudas; en tales situaciones, los números positivos representaban activos. El primer uso conocido de tales números se debe a Brahmagupta, hacia 628; él da también las reglas de las cuatro operaciones para los números negativos. Bhaskara indica que la raíz cuadrada de un número positivo tiene dos valores, uno positivo y otro negativo. Evita la dificultad de la raíz cuadrada de un número negativo, pero afirma que no hay ninguna raíz cuadrada de un número negativo porque estos números no son un cuadrado. No se da ningún tipo de definiciones, axiomas o teoremas.

Los indios no aceptaron los números negativos de manera incondicional. Bhaskara, al dar 50 y -5 como las dos soluciones de un problema, dice: «El segundo valor no debe ser tenido en cuenta en este caso, ya que es inadecuado; la gente no acepta las soluciones negativas.» Sin embargo, los números negativos fueron ganando aceptación lentamente.

Los indios dieron otro gran paso en aritmética al afrontar la cuestión de los números irracionales; es decir, comenzaron a operar con estos números con métodos correctos, lo cual, pese a que no fue demostrado generalmente por ellos, al menos permitió la obtención de conclusiones útiles. Por ejemplo, Bhaskara dice: «Llamemos la suma de dos irracionales al mayor número irracional, y dos veces su producto al menor de ellos. La suma y la diferencia de ellos se efectúa como si fueran números enteros.» Muestra entonces cómo sumarlos: dados los irracionales $\sqrt{3}$ y $\sqrt{12}$,

$$\sqrt{3} + \sqrt{12} = \sqrt{(3 + 12) + 2\sqrt{3 \cdot 12}} = \sqrt{27} = 3\sqrt{3}.$$

El principio general en nuestra notación es

$$\sqrt{a} + \sqrt{b} = \sqrt{(a + b) + 2\sqrt{ab}}. \qquad (1)$$

Destacaríamos la frase «efectuado como números enteros» en la

fórmula anterior. Los irracionales eran tratados como si tuvieran las mismas propiedades que los enteros. Así, si tenemos los enteros c y d podríamos escribir:

$$c + d = \sqrt{(c + d)^2} = \sqrt{c^2 + d^2 + 2cd}. \qquad (2)$$

Ahora si $c = \sqrt{a}$ y $d = \sqrt{b}$, (2) coincide con (1).

Bhaskara da también la regla siguiente para la suma de dos irracionales: «La raíz del cociente del mayor irracional dividida por el menor, aumentada en una unidad; la suma elevada al cuadrado y multiplicada por la menor cantidad irracional es igual a la suma de las dos raíces irracionales.» Esto significa, por ejemplo

$$\sqrt{3} + \sqrt{12} = \sqrt{\left(\sqrt{\frac{12}{3}} + 1\right)^2 \cdot 3},$$

que da $3\sqrt{3}$. Da también reglas para la multiplicación, división y raíz cuadrada de expresiones irracionales.

Los indios eran menos sofisticados que los griegos a la hora de detectar las dificultades lógicas implícitas en el concepto de número irracional. Su interés en el cálculo les hizo pasar por encima de consideraciones filosóficas, o cuestiones que los griegos creían eran fundamentales. No obstante, al aplicar alegremente a los irracionales métodos semejantes a los usados con los racionales ayudaron al progreso de las matemáticas. Además, toda su aritmética fue completamente independiente de su geometría.

Los indios hicieron también algún progreso en álgebra. Usaron abreviaturas de palabras y algunos símbolos para describir las operaciones. Como en el caso de Diofanto, no había ningún símbolo para la adición; una tilde sobre el sustraendo indicaba sustracción; otras operaciones se designaban con palabras clave o abreviaturas; por ejemplo *ka*, de la palabra *karama*, indicaba la raíz cuadrada. Para las incógnitas, cuando había más de una, tenían palabras que denotaban colores. La primera se llamaba la incógnita y las restantes, negro, azul, amarillo y así sucesivamente. Este simbolismo, aunque no era exhaustivo, era suficiente para que se pueda clasificar el álgebra hindú como cuasisimbólica, y en realidad lo era más que el álgebra sincopada de Diofanto. Los problemas y sus soluciones se escribían en este estilo cuasisimbólico. Sólo se daban los pasos y no iban acompañados de justificaciones ni demostraciones.

Los indios sabían que las ecuaciones cuadráticas tenían dos raíces e incluían las negativas y las irracionales. Los tres tipos de ecuaciones cuadráticas $ax^2 + bx = c$, $ax^2 = bx + c$, $ax^2 + c = bx$ con a, b, c positivos, estudiadas por Diofanto de manera independiente, fueron tratadas como un solo caso $px^2 + qx + r = 0$, porque admitían que algunos coeficientes podían ser negativos. Usaban el método de completar un cuadrado, que por supuesto no era nuevo para ellos. Como no admitían las raíces cuadradas de los números negativos, no resolvieron todas las ecuaciones de este tipo. Mahavira resuelve también $x/4 + 2\sqrt{x} + 15 = 0$, que proviene de un problema enunciado verbalmente.

En las ecuaciones indeterminadas avanzaron más allá de Diofanto. Estas ecuaciones surgieron en problemas de astronomía; las soluciones mostraban cuándo ciertas constelaciones aparecerían en el firmamento. Los indios consideraban todas las soluciones enteras mientras que Diofanto tomaba una solución racional. El procedimiento para obtener las soluciones enteras de $ax \pm by = c$ donde a, b y c son números enteros positivos fue introducido por Aryabhata y mejorado por sus sucesores. Es el mismo que el utilizado en la actualidad. Consideremos $ax + by = c$. Si a y b tienen un factor común m que no divide a c, no existe ninguna solución entera ya que el primer miembro es divisible por m mientras que el segundo, no. Si a, b y c tienen un factor común pueden existir soluciones, y, a la vista de la observación precedente, es suficiente considerar el caso en que a y b son primos entre sí. Ahora, el algoritmo de Euclides para el cálculo del máximo común divisor de dos enteros a y b con $a > b$ comienza dividiendo a entre b con lo que $a = a_1b + r$, donde a_1 es el cociente y r, el resto. Por tanto, $a/b = a_1 + r/b$. Esto puede expresarse como:

$$a/b = a_1 + 1(b/r). \qquad (3)$$

El segundo paso en el algoritmo de Euclides consiste en dividir b por r con lo que $b = a_2r + r_1$ o bien $b/r = a_2 + r_1/r$. Si sustituimos el valor de b/r en (3) podemos escribir

$$\frac{a}{b} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{r/r_1}}. \qquad (4)$$

Siguiendo con el algoritmo de Euclides llegamos a lo que se denomina una fracción continua

$$\frac{a}{b} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}},$$

que se escribe también como

$$\frac{a}{b} = a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\dots$$

El proceso se aplica también cuando $a < b$. En este caso, a_1 es cero y entonces, el proceso se continúa como antes. Si a y b son enteros, la fracción continua es finita.

Las fracciones obtenidas al detenerse en el primero, segundo, tercero y en general n-ésimo cociente, reciben el nombre de primer, segundo, tercer, n-ésimo convergente, respectivamente. Como en el caso en que a y b son enteros la fracción continua finaliza, existe un convergente que precede inmediatamente a la expresión exacta de a/b. Si p/q es el valor de este convergente, se puede probar que:

$$aq - bp = \pm 1.$$

Consideremos el valor positivo. Volviendo de nuevo a nuestra ecuación indeterminada original, y como $aq - bp = 1$ podemos escribir:

$$ax + by = c(aq - bp)$$

y reordenando los términos, tenemos

$$\frac{cq - x}{b} = \frac{y + cp}{a}.$$

Si t representa cada una de estas fracciones tenemos:

$$x = cq - bt \quad \text{e} \quad y = at - cp. \tag{5}$$

Podemos asignar ahora valores enteros a t, y como las restantes

cantidades son enteras, obtenemos valores enteros para x e y. En los casos en que $aq - bp = -1$ o cuando la ecuación original es $ax - by = c$, se dan las pequeñas modificaciones que deben realizarse. Brahmagupta da la solución (5) si bien, naturalmente, en función de letras generales a, b, p y q.

Los indios trabajaron también con ecuaciones cuadráticas indeterminadas. Resolvieron el tipo:

$$y^2 = ax^2 + 1, \qquad a \text{ no cuadrado perfecto}$$

y reconocieron que este tipo era fundamental para estudiar la ecuación

$$cy^2 = ax^2 + b.$$

Los métodos utilizados son excesivamente especializados para ser objeto de consideración aquí.

Es destacable que encontraron placer en varios problemas matemáticos y los enunciaron de manera ingeniosa o en verso, o en algún contexto histórico, para agradar y atraer a la gente. La razón original para actuar así pudo haber sido un intento de ayudar la memoria, ya que la vieja práctica del brahmán era confiar en la memoria y escribir las cosas después.

El álgebra se aplicó a los problemas habituales del comercio: cálculo del interés, descuento, división de los beneficios de un socio y la asignación de porciones en una herencia; pero la astronomía fue la aplicación más importante.

3. Geometría y trigonometría indias durante el período 200-1200

Durante este período, la geometría no hizo avances notables; consistía en fórmulas (correctas e incorrectas) para el cálculo de áreas y volúmenes. Muchas de ellas, como la fórmula de Herón para el área de un triángulo y el teorema de Ptolomeo, proceden de los griegos alejandrinos. Algunas veces los indios tenían conciencia de cuándo una fórmula era sólo aproximadamente correcta y otras veces, no. Sus valores para π eran por lo general incorrectos; utilizaban corrientemente $\sqrt{10}$, aunque el valor 3,1416 aparece algunas veces. Para el área de un cuadrilátero dieron la fórmula $\sqrt{(s-a)(s-b)(s-c)(s-d)}$, donde s es el semiperímetro y a, b, c y d son los lados del cuadrilátero,

una fórmula que sólo es correcta para cuadriláteros que pueden inscribirse en un círculo. No presentaron ninguna demostración geométrica; en general, se preocuparon poco por la geometría.

En trigonometría los indios hicieron algunos avances de poca consideración. Ptolomeo había usado las cuerdas de arcos, calculadas sobre la base de que el diámetro de un círculo estaba dividido en 120 unidades. Varahamihira utilizó 120 unidades para el radio. Por tanto, la tabla de cuerdas de Ptolomeo se convirtió para ellos en una tabla de semicuerdas, pero todavía asociadas a todo el arco. Aryabhata hizo entonces dos cambios. Primero, asoció la semicuerda con la mitad del arco de la cuerda completa, este concepto hindú de seno fue usado por todos los matemáticos posteriores. En segundo lugar, introdujo un radio de 3438 unidades. Este número se obtiene asignando $360 \cdot 60$ unidades (el número de minutos) a la circunferencia de un círculo y usando la fórmula $C = 2\pi r$, con π aproximado por 3,14. De esta manera, en el esquema de Aryabhata el seno de un arco de 30°, es decir, la longitud de la semicuerda correspondiente a un ángulo de 30°, era 1719. Aunque utilizaban el equivalente a nuestro coseno, usaban con más frecuencia el seno del arco complementario. Usaban también la noción de «seno verso» ó 1–coseno.

Como el radio de un círculo contenía ahora 3438 unidades, los valores de las cuerdas de Ptolomeo no eran los más adecuados y calcularon de nuevo una tabla de semicuerdas, partiendo de la base de que la semicuerda que corresponde a un arco de 90° es 3438 y la semicuerda que corresponde a un arco de 30° es 1719. Entonces, usando identidades trigonométricas tales como las establecidas por Ptolomeo, fueron capaces de calcular las semicuerdas a intervalos de 3°45′. Este ángulo resulta de dividir cada cuadrante de 90° en 24 partes. Es digno de mención que usaron las identidades en forma algebraica, al contrario de los argumentos de Ptolomeo, e hicieron cálculos aritméticos sobre las relaciones algebraicas. Su práctica era, en principio, semejante a la nuestra.

La motivación de la trigonometría fue la Astronomía, de la que la trigonometría era prácticamente un subproducto. Los trabajos astronómicos típicos incluyen el *Surya Siddhanta* (Sistema del Sol, siglo IV) y el *Aryabhatiya* de Arabhata (siglo VI). El trabajo más importante fue el *Siddhanta Siromani* (Diadema de un sistema astronómico) escrito por Bhaskara en el 1150. Dos capítulos de este trabajo se titulan *Lilavati* (Lo hermoso) y *Vija-ganita* (Extracción de raíces), y estaban dedicados a la aritmética y el álgebra.

Aunque la astronomía constituyó un interés primordial en el período que sigue al año 200, los indios no hicieron grandes progresos en este campo. Se fijaron en la actividad helenística de menor importancia en astronomía aritmética (de origen babilonio), que predice las posiciones planetarias y lunares por extrapolación a partir de los datos de las observaciones. Incluso las palabras hindúes de centro, minuto y otros términos eran exactamente una transliteración de las correspondientes palabras griegas. Los hindúes se interesaron levemente en la teoría geométrica de deferente y epiciclo, si bien enseñaban la esfericidad de la Tierra.

Alrededor del año 1200 declinó la actividad científica en la India y cesó el progreso matemático. Después de que los británicos conquistaran la India en el siglo XVIII, algunos sabios de la India viajaron a Inglaterra para estudiar y a su regreso iniciaron alguna investigación. No obstante, esta actividad moderna forma parte de las matemáticas europeas.

Como indica nuestro resumen, los hindúes estuvieron interesados y realizaron aportaciones en actividades aritméticas y computacionales más que en cuestiones de carácter deductivo. Su nombre para las matemáticas era *Ganita*, que significa «la ciencia del cálculo». Evidenciaron métodos muy buenos y grandes facilidades técnicas, pero no hay ninguna prueba de que consideraran algún tipo de demostración. Tenían reglas, pero en apariencia, ningún escrúpulo lógico. Además, no aportaron ningún método general ni ningún punto de vista novedoso en el área de las matemáticas.

Es rigurosamente cierto que los indios no apreciaban la importancia de sus propias aportaciones. Las escasas ideas valiosas que tuvieron, tales como separar los símbolos correspondientes a los números entre 1 y 9, la conversión a la base 10 y los números negativos fueron introducidos por casualidad sin tener conciencia de que eran innovaciones notables. No eran sensibles a los valores de las matemáticas. En el mundo de las ideas, avanzaron, aceptaron e incorporaron las ideas rudimentarias de egipcios y babilonios. El historiador persa al-Biruni (973-1048) dice de ellos: «Sólo puedo comparar su literatura matemática y astronómica... a una mezcla de concha de perla y dátiles verdes o de perlas y estiércol, o de cristales valiosos y piedras corrientes. Ambos tipos de cosas son iguales a sus ojos porque no pueden elevarse a los métodos de una deducción estrictamente científica.»

4. Los árabes

Visto desde la distancia, el papel de los árabes en la historia de las matemáticas fue el de asestar el último golpe a la civilización alejandrina. Antes de comenzar sus conquistas habían sido un pueblo nómada que ocupaba la región de la Arabia moderna. Fueron incitados a la actividad y a la unidad por Mahoma y menos de un siglo después de su muerte, ocurrida el 632, habían conquistado tierras que iban de la India hasta España incluyendo el norte de Africa y el sur de Italia. El año 755 el imperio árabe se escindió en dos reinos independientes; la parte oriental tenía su capital en Bagdad y la occidental, en Córdoba, en España.

Una vez terminadas sus conquistas, los antiguos nómadas pusieron su empeño en construir una civilización y una cultura. Rápidamente, los árabes se interesaron por las artes y las ciencias. Las dos capitales atrajeron a científicos y apoyaron su trabajo, si bien fue Bagdad la que demostró ser la más importante: allí se construyeron una academia, una biblioteca y un observatorio astronómico.

Los recursos culturales al alcance de los árabes fueron considerables. Invitaron a científicos indios a establecerse en Bagdad. Cuando Justiniano cerró la Academia de Platón el año 529, muchos de sus miembros griegos marcharon a Persia, y las enseñanzas griegas que florecieron allí se convirtieron, un siglo más tarde, en parte del mundo árabe. Los árabes establecieron también contactos con los griegos del Imperio Bizantino; de hecho, los califas árabes adquirieron manuscritos griegos a los bizantinos. Egipto, el centro del saber griego durante el período alejandrino, fue conquistado por los árabes, por lo que la ciencia que sobrevivió allí contribuyó a la actividad del imperio árabe. Las escuelas sirias de Antioquía, Emesa y Damasco y la escuela de los cristianos nestorianos de Edesa, que se habían convertido en los mayores depositarios del Cercano Oriente de los trabajos griegos después de la destrucción de Alejandría el año 640, y los monasterios cristianos del Oriente próximo, que también estaban en posesión de estos trabajos, estaban bajo el gobierno de los árabes. De esta manera, los árabes tenían el control, o el acceso, de hombres y cultura del Imperio Bizantino, Egipto, Siria, Persia y las tierras situadas más al Este, incluida la India.

Se habla de matemáticas árabes, pero en un principio eran matemáticas en lengua árabe. La mayoría de los sabios eran griegos, cristianos, persas y judíos. Sin embargo, es cierto que los árabes, tras

finalizar el período de sus conquistas, marcado por el fanatismo religioso, fueron liberales con respecto a otros pueblos y sectas y los infieles pudieron desarrollar sus actividades con entera libertad.

Fundamentalmente, lo que poseían los árabes era el conocimiento griego, adquirido directamente de los manuscritos griegos o de versiones sirias o hebreas. Todos los trabajos que tenían una gran importancia eran accesibles para ellos. Los bizantinos les proporcionaron una copia de los *Elementos* de Euclides alrededor del año 800 y los tradujeron al árabe. La *Sintaxis Matemática* de Ptolomeo fue traducida también al árabe el año 827 y se convirtió en un libro fundamental, casi divino, para los árabes; era conocido como el *Almagesto*, que significa el libro mayor. Tradujeron también el *Tetrabiblos* de Ptolomeo y este libro de astrología fue popular entre ellos. Con el tiempo fueron accesibles en lengua árabe los trabajos de Aristóteles, Apolonio, Arquímedes, Herón y Diofanto y las obras indias. Entonces, los árabes mejoraron las traducciones e introdujeron comentarios. Estas traducciones, algunas de ellas conservadas, pudieron encontrarse más tarde en Europa, cuando los originales griegos ya se habían perdido. Hasta el 1300 la civilización árabe fue dinámica y su ciencia se difundio ampliamente.

5. Aritmética y álgebra árabes

Cuando los árabes eran todavía nómadas tenían palabras para los números, pero no disponían de ningún símbolo. Tomaron y mejoraron los símbolos numéricos de los indios y su idea de la notación posicional. Usaban estos símbolos numéricos para los números enteros y las fracciones corrientes (añadiendo una barra al esquema hindú) en sus textos matemáticos y numerales alfabéticos árabes, a partir de la idea griega, para los textos astronómicos. Para la astronomía usaban las fracciones sexagesimales, igual que Ptolomeo.

Del mismo modo que los indios, los árabes trabajaron libremente con los irracionales. De hecho, Omar Khayyam (1048?-1122) y Nasîr-Eddin (1201-1274) afirman claramente que toda razón de magnitudes, tanto conmensurables como inconmensurables, puede ser considerada como un número, aseveración que Newton se vio obligado a reafirmar en su *Aritmética Universal* de 1707. Los árabes consideraron las operaciones con números irracionales que habían introducido los indios, y transformaciones tales como $\sqrt{a^2b} = a\sqrt{b}$ y $\sqrt{ab} = \sqrt{a}\,\sqrt{b}$ se convirtieron en habituales.

En aritmética, los árabes dieron un paso atrás: aunque estaban familiarizados con los números negativos y las reglas de las operaciones con ellos a través de los trabajos de los indios, los rechazaron.

Al álgebra contribuyeron antes de nada con el nombre. La palabra «álgebra» viene de un libro escrito el 830 por el astrónomo Mohammed ibn Musa al-Khowârizmî (sobre el 825), titulado *Al-jabr w'al muqâbala*. La palabra *al-jabr* que en este contexto significa «restauración», restaura el equilibrio en una ecuación al colocar en un miembro de la misma un término que ha sido eliminado del otro; por ejemplo, si -7 se suprime de $x^2 - 7 = 3$, el equilibrio se restaura escribiendo $x^2 = 7 + 3$. *Al' muqâbala* significa «simplificación», en el sentido de que, por ejemplo, se pueden combinar $3x$ y $4x$ y obtener $7x$, o bien suprimir términos iguales en miembros distintos de una ecuación. *Al-jabr* significa también «componedor de huesos», es decir, un restaurador de huesos rotos. Cuando los moros llevaron la palabra a España se convirtió en *algebrista* y significaba «componedor de huesos». En algún tiempo no era raro en España ver un cartel con la inscripción «Algebrista y Sangrador» [1] a la entrada de una barbería, ya que en aquel tiempo, e incluso en siglos posteriores, los barberos administraban tratamientos médicos sencillos. En el siglo XVI en Italia, álgebra significaba el arte de componer huesos rotos. Cuando el libro de al-Khowârizmî fue traducido por primera vez al latín en el siglo XII, se tituló *Ludus algebrae et almucgrabalaeque*, aunque se usaron también otros títulos. El nombre fue finalmente resumido como álgebra.

El álgebra de al-Khowârizmî está basada en el trabajo de Brahmagupta, pero muestra también influencias babilonias y griegas. Al-Khowârizmî ejecuta algunas operaciones exactamente igual que Diofanto. Por ejemplo, en ecuaciones con varias incógnitas, las reduce a una indeterminada y a continuación las resuelve. Diofanto coloca una indeterminada a continuación de otra para escribir s^2 y lo mismo hace al-Khowârizmî. Este llama «potencia» al cuadrado de la incógnita, que es una palabra de Diofanto. Utiliza también, igual que Diofanto, nombres especiales para las potencias de la indeterminada. Llama a esta última la «cosa» o la «raíz» (de una planta), y de ahí procede nuestro término raíz. Al-Karkhî de Bagdad (fallecido sobre el 1029) que escribió un texto árabe de álgebra superior en los primeros años del siglo XI, sigue realmente a los griegos y en especial a Diofanto. Sin

---

[1] *N. T.* En español en el original.

embargo, los árabes no usaron ninguna clase de simbolismo; su álgebra es completamente retórica y a este respecto representa un retroceso si se compara con la de los indios e incluso con la de Diofanto.

En su Algebra, al-Khowârizmî da el producto de $(x \pm a)$ e $(y \pm b)$. Muestra cómo añadir y sustraer términos de expresiones de la forma $ax^2 + bx + c$. Resuelve ecuaciones lineales y cuadráticas, pero considera las seis formas distintas, tales como $ax^2 = bx$, $ax^2 = c$, $ax^2 + c = bx$, $ax^2 + bx = c$ y $ax^2 = bx + c$, con a, b, c siempre positivos. Evita los números negativos en solitario y la sustracción de cantidades mayores que el minuendo. En este procedimiento de considerar formas separadas, al-Khowârizmî sigue la línea de Diofanto. Al-Khowârizmî reconoce que una ecuación cuadrática puede tener dos raíces, pero da solamente las que son reales y positivas, que pueden ser irracionales. Algunos autores dan raíces positivas y raíces negativas.

Un ejemplo de un problema cuadrático planteado por al-Khowârizmî es el siguiente: «Un cuadrado y diez de sus raíces son iguales a treinta y nueve unidades, es decir, si sumamos diez raíces a un cuadrado, la suma es igual a treinta y nueve.» Da la solución así: «Tomemos la mitad del número de raíces, esto es, en este caso, cinco, y multipliquemos esta cantidad por sí misma y el resultado es veinticinco. Añadámosla a treinta y nueve, lo que da sesenta y cuatro; tomemos su raíz cuadrada, u ocho, y restémosle la mitad del número de raíces, precisamente cinco, y queda un resto de tres. Esta es la raíz.» La solución dada sigue exactamente el proceso llamado «completar un cuadrado».

Pese a que los árabes dan soluciones algebraicas de las ecuaciones cuadráticas, explican o justifican sus procesos geométricamente. Sin duda, estaban influidos por la confianza que tenían los griegos en el álgebra geométrica: a la vez que aritmetizaban el problema, debían pensar que la demostración había que hacerla con métodos geométricos. Así, para resolver la ecuación $x^2 + 10x = 39$, al-Khowârizmî da el siguiente razonamiento geométrico: sea AB (fig. 9.1) el segmento que representa el valor de la incognita x y construyamos el cuadrado $ABCD$. Prolonguemos DA hasta H y DC hasta F de manera que $AH = CF = 5$, que es la mitad del coeficiente de x. Completemos el cuadrado sobre DH y DF. Entonces, las áreas I, II y III son x^2, $5x$ y $5x$ respectivamente. La suma de las tres es el primer miembro de la ecuación. Añadimos ahora a ambos miembros el área IV, que es 25.

Luego, el cuadrado completo tiene área 39 + 25 ó 64 y su lado debe valer 8. Así pues, *AB* o *AD* es 8 − 5 ó 3. Este es el valor de *x*. El argumento geométrico se basa en la proposición 4 del libro II de los *Elementos*.

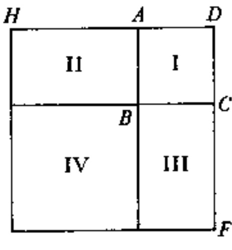


FIGURA 9.1

Los árabes resolvieron algunas ecuaciones cúbicas algebraicamente y dieron una justificación geométrica, tal como hemos ilustrado en el caso de las ecuaciones cuadráticas. Esto es lo que hicieron, por ejemplo, Tâbit ibn Qorra (836-901), un pagano de Bagdad, que fue también físico, filósofo y astrónomo, y el egipcio al-Hassan ibn al-Haitham, conocido generalmente como Alhazen (sobre 965-1039). Omar Khayyam creía que la ecuación cúbica general debería resolverse sólo geométricamente, usando secciones cónicas. Vamos a ilustrar el método usado en su *Algebra* (hacia 1079) para resolver algunos tipos de tales ecuaciones considerando uno de los casos más sencillos estudiados por él: $x^3 + Bx = C$, donde B y C son positivos.

Khayyam escribe la ecuación como $x^3 + b^2x = b^2c$ donde $b^2 = B$ y $b^2c = C$. Construye entonces una parábola (fig. 9.2) de parámetro b. Esta cantidad, por supuesto, fija la parábola, y aunque la curva no puede construirse con el uso de la regla y el compás, se pueden dibujar

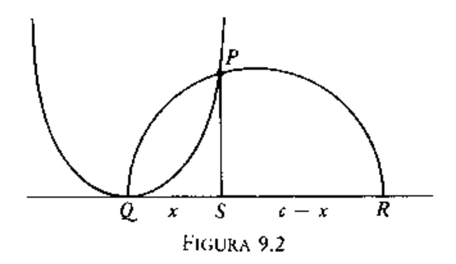


FIGURA 9.2

cuantos puntos de la misma se deseen. Traza a continuación el semicírculo de diámetro QR, que tiene longitud c. La intersección P de la parábola y el semicírculo determina la perpendicular PS, y QS es la solución de la ecuación cúbica.

La demostración de Khayyam es estrictamente sintética. A partir de la propiedad geométrica de la parábola dada por Apolonio (o como podemos ver a partir de la ecuación $x^2 = by$)

$$x^2 = b \cdot PS \tag{6}$$

o bien

$$\frac{b}{x} = \frac{x}{PS}. \tag{7}$$

Consideremos ahora el triángulo rectángulo QPR. La altura PS es media proporcional entre QS y SR. Por tanto:

$$\frac{x}{PS} = \frac{PS}{c - x}. \tag{8}$$

De (7) y (8) obtenemos

$$\frac{b}{x} = \frac{PS}{c - x}. \tag{9}$$

Pero por (7)

$$PS = \frac{x^2}{b}.$$

Si sustituimos este valor de PS en (9) vemos que x satisface la ecuación $x^3 + b^2x = b^2c$.

Khayyam resolvió también ecuaciones del tipo $x^3 + ax^2 = c^3$, cuyas raíces están determinadas por la intersección de una hipérbola y una parábola, y del tipo $x^3 \pm ax^2 + b^2x = b^2c$, las raíces de la cual se determinan mediante la intersección de una elipse y una hipérbola. Resolvió asimismo una ecuación de cuarto grado: $(100 - x^2)(10 - x^2) = 8100$, cuyas raíces están determinadas por la intersección de una hipérbola y un círculo. Da solamente raíces positivas.

La resolución de ecuaciones cúbicas con el uso de intersecciones de cónicas es el mayor avance hecho por los árabes en álgebra. Las matemáticas son del mismo tipo que el álgebra geométrica de los griegos, aunque utilicen secciones cónicas. El objetivo sería dar una respuesta aritmética, pero los árabes la podían obtener solamente midiendo la longitud final que representa x. En este trabajo la influencia de las geometría griega es evidente.

Los árabes resolvieron también ecuaciones indeterminadas de segundo y tercer grado. Un par de autores se dedicaron a estudiar la ecuación $x^3 + y^3 = z^3$, pero no pudo ser resuelta completamente. Dieron también las sumas de las potencias primera, segunda, tercera y cuarta de los n primeros números naturales.

6. La geometría y la trigonometría árabes

La geometría árabe estuvo influida principalmente por Euclides, Arquímedes y Herón. Los árabes hicieron comentarios críticos a los *Elementos* de Euclides, lo que resulta sorprendente, ya que pone de manifiesto unas apreciaciones de rigor que contrastan con su indiferencia habitual por el mismo en álgebra. Estos comentarios incluyen trabajos sobre el axioma de las paralelas, que consideraremos más adelante (cap. 36). Estos comentarios son más valiosos por la información que suministran acerca de los manuscritos griegos que los árabes tenían a su disposición desde que habían sido extraviados que por el hecho de que pudieran ofrecer algún resultado nuevo o alguna demostración diferente. Un problema nuevo, que se hizo popular en Europa durante el Renacimiento, fue investigado por el persa Abû'l-Wefâ o Albuzjani (940-998); construcciones con una recta y un círculo fijo (es decir, un compás con abertura fija).

Los árabes hicieron algún pequeño progreso en trigonometría. La suya, igual que en el caso de los indios, es más aritmética que geométrica (como en Hiparco y Ptolomeo). Así, para calcular valores de algún coseno a partir de valores del seno usaban una identidad como $\text{sen}^2 A + \cos^2 A = 1$ y transformaciones algebraicas. Igual que los indios, utilizaban senos de arcos en vez de cuerdas de arcos dobles, pese a que (como en los trabajos indios) el número de unidades en el seno o semicuerda depende del número de unidades tomadas en el radio. Tâbit ibn Qorra y el astrónomo al-Battânî (sobre 858-929) introdujeron este uso de senos entre los árabes.

Los astrónomos árabes introdujeron lo que llamamos tangente y cotangente, pero como líneas que contenían un número determinado de unidades, exactamente como el seno de un arco era una longitud que contenía una cantidad dada de unidades. Estas dos razones se pueden encontrar en el trabajo de al-Battânî. Abû'l-Wefâ introdujo la secante y la cosecante como longitudes en un trabajo de astronomía. Calculó también tablas de senos y tangentes para intervalos de 10′ de ángulo. Al-Bîrunî dio el teorema del seno para triángulos planos y una demostración del mismo.

La sistematización de la trigonometría plana y esférica en un trabajo independiente de la astronomía la proporcionó Nasîr-Eddin en su *Tratado del Cuadrilátero*. Este trabajo contiene seis fórmulas fundamentales para la resolución de triángulos rectángulos esféricos y muestra cómo resolver triángulos más generales por el método que denominamos del triángulo polar. Por desgracia, los europeos no conocieron el trabajo de Nasîr hasta aproximadamente el año 1450; hasta entonces la trigonometría permaneció, tal como había sido concebida, como un apéndice de la astronomía, tanto en los textos como en las aplicaciones.

El esfuerzo científico árabe, si bien no fue original, fue amplio; sin embargo, nosotros no podemos hacer otra cosa que destacar del mismo la continuación de las líneas propuestas por los griegos. Contrariamente a los indios, los árabes tomaron como punto de partida la astronomía de Ptolomeo. La astronomía fue cobrando cada vez más importancia, ya que las horas de las oraciones debían ser conocidas con absoluta precisión, y porque los árabes debían rezar de cara a La Meca desde cualquier punto de su vasto imperio. Las tablas astronómicas fueron aumentadas; los instrumentos, mejorados, y se construyeron y utilizaron observatorios. Igual que en la India, prácticamente todos los matemáticos eran principalmente astrónomos. La astrología jugó también un importante papel a la hora de estimular los trabajos en astronomía y, en consecuencia, en matemáticas.

Otra ciencia estudiada por los árabes fue la óptica. Alhazen, que fue físico además de matemático, escribió el gran tratado *Kitab al-manazer* o *Compendio de Optica*, que ejerció una gran influencia. En él se establece la ley de la reflexión completa, incluyendo el hecho de que el rayo incidente, el reflejado y la normal a la superficie de reflexión están todos contenidos en un mismo plano. Pero, igual que Ptolomeo, no tuvo éxito a la hora de hallar la ley del ángulo de refracción, pese a que dedicó a ello grandes esfuerzos y experimenta-

ciones. Habló de los espejos esféricos y parabólicos, las lentes, la cámara oscura y la visión. La óptica fue un tema favorito de los árabes porque se presta a proyectos ocultos y místicos. Pese a ello, no aportaron ninguna idea original de importancia.

El uso de las matemáticas por parte de los árabes se centraba en los campos que hemos indicado anteriormente. La astronomía, la astrología, la óptica y la medicina (a través de la astrología) las necesitaban, si bien partes del álgebra, como dijo uno de los matemáticos árabes, eran «más necesarias en cuestiones de distribución, herencias, sociedades, medidas de tierras...». Los árabes estudiaron matemáticas para el resto de las pocas ciencias que cultivaron, y no por sí mismas. No estudiaron las ciencias por su valor intrínseco. No se interesaron en el objetivo de los griegos de comprender el plan matemático de la Naturaleza o para entender los caminos de Dios, como en la Europa medieval. El objetivo árabe, nuevo en la historia de la ciencia, era dominar la Naturaleza. Ellos pensaban que lograrían este poder a través de la alquimia, la magia y la astrología, que constituyeron una gran parte de su esfuerzo científico. Este objetivo fue abandonado más tarde por mentes más críticas, que distinguirían entre ciencia y pseudociencia y fueron más profundos en su estudio.

Los árabes no produjeron ningún avance significativo en matemáticas. Lo que hicieron fue absorber las matemáticas griegas e indias, conservarlas y finalmente, a través de los acontecimientos que hemos contemplado hasta ahora, transmitirlas a Europa. La actividad árabe alcanzó su cumbre alrededor del año 1000. Entre los años 1100 y 1300 los ataques cristianos en las Cruzadas debilitaron a los árabes orientales. Como consecuencia de ello, su territorio fue invadido y conquistado por los mongoles; después de 1258, el califato de Bagdad dejó de existir. Destrucciones adicionales a cargo de los tártaros bajo Tamerlán acabaron de borrar esta civilización árabe, aunque se desarrolló algún atisbo de actividad matemática después de la invasión tártara. En España, los árabes fueron constantemente atacados y finalmente derrotados en el año 1492 por los cristianos; esto terminó con la actividad matemática y científica en ese país.

7. La matemática alrededor del 1300

Si bien la labor matemática de los indios y árabes no fue brillante, produjo algunos cambios en el contenido y en el carácter de las

matemáticas que fueron importantes para abordar su futuro. La notación posicional en base 10 (que usa símbolos especiales para los números de 1 a 9 y el cero como número), la introducción de los números negativos y el libre uso de los irracionales como números no sólo amplió considerablemente la aritmética sino que allanó el camino para un álgebra más trascendente, un álgebra en la que las letras y las operaciones se podrían aplicar a una clase más amplia de números.

Ambos pueblos trabajaron con ecuaciones, determinadas e indeterminadas, con una base más aritmética que geométrica. Si bien el álgebra, tal como la iniciaron egipcios y babilonios, estaba fundamentada aritméticamente, los griegos la habían desvirtuado al requerir una base geométrica. Por lo tanto, los trabajos indios y árabes no sólo recondujeron el álgebra a sus propios orígenes sino que incluso la hicieron avanzar por varios caminos. Los indios aumentaron el simbolismo e hicieron progresos en el campo de las ecuaciones indeterminadas, mientras que los árabes se aventuraron en el problema de las ecuaciones de tercer grado, aunque los trabajos de Khayyam todavía estaban vinculados a la geometría.

La geometría euclídea no hizo ningún avance, pero sí lo hizo la trigonometría. La introducción del seno o semicuerda demostró ser un avance técnico. La técnica aritmética o algebraica para el manejo de identidades y para el cálculo en trigonometría fue un paso definitivo, y la segregación del conocimiento trigonométrico de la astronomía puso de manifiesto una ciencia más ampliamente aplicable.

Hubo dos hechos que fueron significativos para el reconocimiento futuro de que el álgebra podía coexistir con la geometría. La aceptación de los números irracionales hizo posible asignar valores numéricos a todos los segmentos lineales y a las figuras de dos o tres dimensiones; es decir, expresar longitudes, áreas y volúmenes mediante números. Además, la práctica de los árabes de resolver ecuaciones algebraicas al mismo tiempo que justificaban el proceso a través de una representación geométrica exhibía el paralelismo entre las dos materias. El desarrollo más completo de este paralelismo se dejó a la geometría analítica.

Quizá más interesante sea el concepto autocontradictorio de las matemáticas que tenían indios y árabes. Ambos trabajaron con entera libertad en aritmética y álgebra y, sin embargo, no se preocuparon en ninguna medida de la noción de demostración. No resulta sorprendente que tanto egipcios como babilonios se contentaran con su escasa aritmética y algunas reglas geométricas de base empírica; ésta es

una base natural para casi todo el conocimiento humano. Pero los indios y los árabes estaban al tanto del concepto totalmente nuevo de demostración matemática instaurado por los griegos. El comportamiento hindú puede tener alguna justificación; pese a que poseían claramente algún conocimiento de los trabajos de los griegos clásicos, les prestaron muy poca atención y siguieron fundamentalmente el tratamiento greco-alejandrino de la aritmética y el álgebra. Incluso su preferencia por un tipo de matemáticas sobre otras levanta una controversia. Pero los árabes conocían completamente la geometría griega e incluso hicieron estudios críticos de Euclides y otros autores griegos. Por otra parte, las condiciones para continuar el estudio de la ciencia pura eran favorables en un período que duró varios siglos, por lo que la presión para producir resultados prácticos y útiles no tenía que dar lugar necesariamente a que los matemáticos sacrificaran la demostración por la utilidad inmediata. ¿Por qué, pues, los dos pueblos han tratado las dos áreas de las matemáticas de forma tan distinta a los griegos?

Hay muchas respuestas posibles. Ambas civilizaciones carecieron por completo de espíritu crítico, a pesar de los comentarios árabes de Euclides. Luego, es posible que se contentaran con tomar las matemáticas tal como las encontraron; es decir, la geometría era deductiva, pero la aritmética y el álgebra eran empíricas o heurísticas. Una segunda posibilidad es que tanto un pueblo como el otro —más propiamente los árabes— admitieron que los métodos y temas de la geometría eran completamente diferentes de los de la aritmética y el álgebra pero no veían la manera de encontrar un fundamento lógico para la aritmética. Un hecho que parece apoyar esta teoría es que los árabes, al menos, explicaban su solución de las ecuaciones cuadráticas con argumentos geométricos.

Hay otras explicaciones posibles. Indios y árabes apoyaron la aritmética, el álgebra y la formulación algebraica de las relaciones trigonométricas, así como las operaciones con ellas. Esta predisposición puede provocar una mentalidad distinta o reflejar una respuesta a las necesidades de las civilizaciones. Las dos que estamos considerando estaban orientadas a la práctica y, como hemos tenido ya ocasión de observar al hablar de los griegos alejandrinos, las necesidades prácticas exigían resultados cuantitativos, que eran facilitados por la aritmética y el álgebra. Una pequeña evidencia que puede favorecer la tesis de que había una diferencia de mentalidad respecto a los griegos clásicos es la reacción de los europeos ante una herencia matemática

muy parecida a la que habían recibido indios y árabes. Como veremos, los europeos estuvieron mucho más preocupados por la situación tan dispar que presentaban la aritmética y la geometría.

En ausencia de un estudio exhaustivo y definitivo podemos adoptar la postura de admitir que los indios y árabes eran conscientes de la situación precaria en que se encontraban la aritmética y el álgebra, pero tuvieron la audacia, reforzada por necesidades prácticas, de desarrollar estas ramas. Si bien es indudable que no apreciaban lo que estaban realizando, tomaron el único camino de innovación matemática posible. Las nuevas ideas pueden llegar solamente a través de la búsqueda libre y audaz del discernimiento heurístico e intuitivo. La justificación lógica y las medidas correctoras, que se necesitarían más tarde, pueden plantearse solamente cuando hay algo que fundamentar lógicamente. Los indios y árabes tuvieron el atrevimiento de llevar de nuevo la aritmética y el álgebra al lugar que ocupaban anteriormente y situarlas casi a la misma altura que la geometría.

Se había llegado así a dos tradiciones o conceptos de matemáticas independientes entre sí: por una parte, el cuerpo de conocimiento lógico y deductivo establecido por los griegos, que servía para el ambicioso propósito de comprender la Naturaleza; y por otra, las fundamentadas empíricamente y orientadas a la práctica, creadas por los egipcios y babilonios y resucitadas por los greco-alejandrinos y prolongadas por los indios y árabes. Una favorece la geometría y la otra, la aritmética y el álgebra. Ambas tradiciones y ambos objetivos fueron continuados y llevados a cabo.

Bibliografía

Ball, W. W. R.: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, cap. 9.

Berry, Arthur: *A Short History of Astronomy*, Dover (reimpresión), 1961, pp. 76-83.

Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Cajori, Florian: *A History of Mathematics*, Macmillan, 1919, pp. 83-112.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1894, Johnson Reprint Corp., 1965, vol. 1, caps. 28-30, 32-37.

Coolidge, Julian L.: *The Mathematics of Great Amateurs*, Dover (reimpresión), 1963, cap. 2.

Datta, B., and A. N. Singh: *History of Hindu Mathematics*, 2 vols. Asia Publishing House (reimpresión), 1962.

Dreyer, J. L. E.: *A History of Astronomy from Thales to Kepler,* Dover (reimpresión), 1953, cap. 11.

Karpinski, L. C.: *Robert of Chester's Latin Translation of the Algebra of al-Khowarizmi,* Macmillan 1915. Hay también versión inglesa.

Kasir, D. S.: *The Algebra of Omar Khayyam,* Columbia University Teachers College thesis, 1931.

O'Leary, De Lacy: *How Greek Sciences Passed to the Arabs,* Routledge and Kegan Paul, 1949.

Pannekoek, A.: *A History of Astronomy,* John Wiley and Sons, 1961, cap. 15.

Scott, J. F.: *A History of Mathematics,* Taylor and Francis, 1958, cap. 5.

Smith, David Eugene: *History of Mathematics,* Dover (reimpresión), 1958, vol. 1, pp. 138-47, 152-92, 283-90.

Struik, D. J.: «Omar Khayyam, Mathematician», *The Mathematics Teacher,* 51, 1958, 280-285.

Capítulo 10
EL PERIODO MEDIEVAL EN EUROPA

> En la mayoría de las ciencias una generación destruye lo que otra ha construido y lo que una ha establecido otra lo deshace. Sólo en matemáticas cada generación añade un piso nuevo a la antigua estructura.
>
> HERMANN HANKEL

1. Los comienzos de la civilización europea

La Europa occidental y central empezó a participar en el desarrollo de las matemáticas cuando la civilización árabe comenzó a declinar. Sin embargo, para familiarizarnos con la situación de la Europa medieval, para saber cómo empezó la civilización europea y para entender las direcciones que tomó, debemos volver, al menos brevemente, a sus comienzos.

En los tiempos en que florecieron babilonios, egipcios, griegos y romanos, la zona que ahora se llama Europa (excepto Italia y Grecia) poseía una civilización primitiva. Las tribus germánicas que vivían en ella no tenían ni una escritura ni una cultura. El historiador romano Tácito (siglo I d. C.) describe esas tribus, aproximadamente en tiempos de Cristo, como honestas, hospitalarias, bebedoras, no amantes de la paz y orgullosas de la lealtad de sus esposas. La cría de ganado, la caza y el cultivo de grano eran las principales ocupaciones. Comenzando en el siglo IV de la era cristiana, los hunos condujeron hacia el Oeste a los godos y a las tribus germánicas que ocupaban la Europa central. En el siglo V los godos conquistaron el Imperio Romano de Occidente.

Aunque partes de Francia e Inglaterra habían adquirido alguna cultura durante la dominación del Imperio Romano, hacia el 500 después de Cristo comenzaron a actuar en Europa nuevas influencias civilizadoras. Incluso antes de la caída del Imperio, la Iglesia Católica estaba organizada y era poderosa. La iglesia convirtió gradualmente a los bárbaros germánicos y godos al cristianismo y comenzó a fundar escuelas; éstas estaban asociadas a monasterios ya existentes que conservaban fragmentos de las culturas griega y romana, y habían estado enseñando a la gente a leer los servicios de la iglesia y los libros sagrados. Un poco más tarde, la necesidad de preparar hombres para desempeñar puestos eclesiásticos motivó el desarrollo de escuelas superiores.

En la última mitad del siglo VIII, algunos dirigentes seglares fundaron más escuelas. En el imperio de Carlogmano, las escuelas fueron organizadas por Alcuino de York (730-804), un inglés que vino a Europa por invitación del mismo Carlomagno. Estas escuelas también estuvieron asociadas a catedrales o monasterios, y enfatizaban la teología cristiana y la música. En realidad, las universidades en Europa se desarrollaron a partir de las escuelas eclesiásticas, con profesores suministrados por las órdenes religiosas, como los franciscanos y los dominicos. Bolonia, la primera universidad, fue fundada en 1088. Las universidades de París, Salerno, Oxford y Cambridge fueron establecidas alrededor de 1200. Por supuesto que, en sus comienzos, difícilmente podían ser consideradas como universidades en el sentido actual. Además, aunque formalmente independientes, estaban esencialmente dedicadas a los intereses de la Iglesia.

2. Los elementos disponibles para la cultura

A medida que la Iglesia extendía su influencia, iba favoreciendo e imponiendo una determinada cultura. El latín era la lengua oficial de la Iglesia y por ello el latín se convirtió en la lengua internacional de Europa y en la lengua de las matemáticas y de la ciencia. Era también la lengua en que se enseñaba en las escuelas europeas hasta bien entrado el siglo XVIII. Resultó inevitable que los europeos buscaran el conocimiento sobre todo en libros latinos, esto es, romanos. Puesto que la matemática romana era insignificante, todo lo que los europeos aprendieron fue un sistema de números muy primitivo y unos pocos

conocimientos de aritmética. También conocieron un poco de la matemática griega a través de algunos traductores.

El principal traductor, cuyas obras fueron ampliamente utilizadas hasta el siglo XII, fue Anicio Manlio Severino Boecio (*c.* 480-524), descendiente de una de las más antiguas familias romanas. Utilizando fuentes griegas, compiló selecciones latinas de tratados elementales sobre aritmética, geometría y astronomía. De los *Elementos,* de Euclides, puede haber traducido lo mismo cinco libros que dos, y éstos constituyeron una parte de su *Geometría.* En este tema dio definiciones y teoremas, pero no demostraciones. También incluyó en este trabajo algún material sobre la geometría de la medición. Algunos resultados son incorrectos y otros, sólo aproximaciones. Curiosamente, la *Geometría* también contenía material sobre los ábacos y las fracciones, estas últimas como preliminar al material sobre astronomía (que no tenemos). Boecio también escribió *Institutis arithmetica,* una traducción de la *Introductio arithmetica* de Nicómaco, aunque omitió algunos de los resultados de éste. Este libro fue la fuente de toda la aritmética que se enseñó en las escuelas durante casi mil años. Finalmente, Boecio tradujo algunas obras de Aristóteles, y escribió una astronomía basada en Ptolomeo y un libro sobre música basado en trabajos de Euclides, Ptolomeo y Nicómaco. Es muy probable que Boecio no entendiera todo lo que traducía. Fue él quien introdujo la palabra «*quadrivium*» para designar a la aritmética, geometría, música y astronomía. Su obra más conocida, todavía leída en la actualidad, son las *Consolaciones de la Filosofía,* que escribió mientras estaba en prisión por supuesta traición (por la que fue decapitado al final).

Otros traductores fueron el romano Aurelio Casiodoro (*c.* 475-570), quien expuso unas pocas partes de las obras griegas en matemáticas y astronomía en su propia pobre versión; Isidoro de Sevilla (*c.* 560-636), quien escribió las *Etimologías,* una obra en veinte libros sobre temas que abarcaban desde las matemáticas hasta la medicina, y el inglés Beda el Venerable (674-735). Estos hombres fueron los eslabones principales entre las matemáticas griegas y los primeros tiempos del mundo medieval.

En todos los problemas que aparecen en los libros escritos por los primeros matemáticos medievales sólo aparecen las cuatro operaciones con enteros. Puesto que, en la práctica, los cálculos se hacían en varios tipos de ábacos, las reglas de estas operaciones estaban especialmente adaptadas para ello. Las fracciones se utilizaban raras veces, y cuando se hacía se utilizaban fracciones romanas con nombres especí-

ficos más que un simbolismo especial; por ejemplo, *uncia* era 1/12, *quincunx* era 5/12, *dodrans* era 9/12. Los números irracionales no aparecían en absoluto. Los buenos calculadores fueron conocidos en la Edad Media como practicantes de una forma de magia, el «Arte Negro».

En el siglo X, Gerbert (1003†), nativo de Auvernia, y más tarde el Papa Silvestre II, contribuyeron a mejorar el estudio de las matemáticas. Sus escritos, sin embargo, se limitaron a la aritmética y geometría elementales.

3. El papel de las matemáticas en Europa en la Alta Edad Media

Aunque los elementos disponibles para la enseñanza de las matemáticas eran escasos, éstas eran relativamente importantes en el *curriculum* de las escuelas medievales, incluso desde muy pronto. El *curriculum* estaba dividido en el *quadrivium* y el *trivium*. El *quadrivium* incluía la aritmética, considerada como la ciencia de los números puros, la música, vista como una aplicación de los números, la geometría, o el estudio de magnitudes tales como longitudes, áreas y volúmenes en reposo, y la astronomía, el estudio de magnitudes en movimiento. El *trivium* cubría la retórica, la dialéctica y la gramática.

Incluso el aprendizaje de las pocas matemáticas que hemos descrito servía para varios propósitos. Después de la época de Gerbert fueron aplicadas para obtener alturas y distancias, para lo que se utilizaban el astrolabio y el espejo como instrumentos de campo. Se esperaba de la clerecía que defendiera la teología y refutara argumentos en contra mediante razonamientos, y las matemáticas eran consideradas como un buen entrenamiento para el razonamiento teológico, de la misma manera que Platón las había considerado como un buen entrenamiento para la filosofía. La Iglesia también abogaba por la enseñanza de las matemáticas por su aplicación para el cumplimiento del calendario y la predicción de las fiestas. En cada monasterio había al menos una persona que podía realizar los cálculos necesarios para ello, y en el curso de este trabajo fueron diseñadas numerosas mejoras en aritmética y en el método de cálculo del calendario.

Otra motivación para el estudio de algunas matemáticas fue la astrología. Esta pseudociencia, que había estado algo en boga en Babilonia, en la Grecia helenística y entre los árabes, era casi univer-

salmente aceptada en la Europa medieval. La doctrina básica de la astrología era, por supuesto, que los cuerpos celestes influían y controlaban los cuerpos humanos y su destino. Para entender las influencias de los cuerpos celestes y para predecir lo que presagiaban acontecimientos celestes especiales, tales como conjunciones y eclipses, eran necesarios algunos conocimientos de astronomía, y por tanto eran indispensables algunas matemáticas.

La astrología fue especialmente importante en la Baja Edad Media. Todas las cortes tenían astrólogos y las universidades tenían profesores de astrología y cursos sobre el tema. Los astrólogos aconsejaban a los príncipes y a los reyes sobre decisiones políticas, campañas militares y asuntos personales. Es curioso que incluso gobernantes instruidos y vinculados al pensamiento griego confiaran en los astrólogos. En el último período medieval y en el Renacimiento, la astrología no sólo se convirtió en una actividad importante, sino que fue considerada como una rama de las matemáticas.

A través de la astrología, se estableció una relación entre las matemáticas y la medicina (cap. 7, sec. 8). Aunque la Iglesia menospreciaba el cuerpo físico como relativamente falto de importancia, los médicos no estaban necesariamente de acuerdo con ello. Puesto que los cuerpos celestes influían presumiblemente en la salud, los médicos estudiaban, por una parte, las relaciones entre los acontecimientos celestes y las constelaciones y, por otra, la salud de los individuos. Se recogían y conservaban los datos de las constelaciones que aparecían en los nacimientos, matrimonios, enfermedades y muertes de miles de personas y se utilizaban para predecir el éxito de los tratamientos médicos. Para realizar todo esto se requería un conocimiento tan amplio de las matemáticas que los médicos tenían que ser personas ilustradas en este campo. De hecho, eran más astrólogos y matemáticos que estudiosos del cuerpo humano.

La aplicación de las matemáticas a la medicina a través de la astrología se extendió todavía más durante la última parte del período medieval. Bolonia tuvo una escuela de medicina y de matemáticas en el siglo XII. Cuando el astrónomo Tycho Brahe acudió a la Universidad de Rostock en 1566, allí no había astrónomos sino astrólogos, alquimistas, matemáticos y médicos. En muchas universidades, los profesores de astrología eran más comunes que los profesores de medicina y de astronomía propiamente dichas. Galileo enseñó astronomía a estudiantes de medicina, pero por su interés para la astrología.

4. El estancamiento en matemáticas

La Alta Edad Media se extiende aproximadamente desde el 400 hasta el 1100: setecientos años, durante los cuales la civilización europea podría haber desarrollado algunas matemáticas. Podría haber encontrado una inmensa ayuda en los trabajos griegos si hubiera seguido las pocas directrices disponibles hacia el vasto conocimiento encerrado en ellas. Aunque durante este período las matemáticas no experimentaron ningún progreso, tampoco se hizo ningún intento serio por construirlas. Las razones pueden ser de interés para quienes desean entender bajo qué circunstancias pueden florecer las matemáticas.

La razón primordial del bajo nivel de las matemáticas era la ausencia de interés por el mundo físico. La Cristiandad, que dominaba Europa, prescribía sus propios fines, valores y modo de vida. Las preocupaciones importantes eran espirituales, de tal manera que los interrogantes sobre la naturaleza que estuvieran estimulados por la curiosidad o por fines prácticos eran considerados como frívolos o sin valor. La Cristiandad, e incluso los últimos filósofos griegos, estoicos, epicúreos y neoplatónicos, resaltaron la elevación de la mente sobre la carne y la materia y la preparación del alma para una vida futura en el cielo. La realidad última era la vida eterna del alma; y la salud del alma se reforzaba mediante el aprendizaje de verdades morales espirituales. Las doctrinas del pecado, del miedo del infierno, de la salvación y del deseo del cielo eran dominantes. Puesto que el estudio de la naturaleza no contribuía a alcanzar tales fines o a prepararse para la vida futura, era rechazado como algo sin valor e incluso herético.

¿De dónde, entonces, sacaron los europeos el conocimiento sobre la naturaleza y plan del universo y del hombre? La respuesta es que todo ese conocimiento fue obtenido del estudio de las Escrituras. Los credos y dogmas de los Padres de la Iglesia, que eran ampliaciones e interpretaciones de las Escrituras, fueron tomados como la suprema autoridad. San Agustín (354-430), hombre muy instruido, y muy influyente en la difusión del neoplatonismo, dijo: «Cualquier conocimiento que el hombre haya adquirido fuera de las Sagradas Escrituras, si es dañino, allí está condenado; si es saludable, allí está contenido.» Esta cita, si bien no es representativa de Agustín, sí lo es de la actitud hacia la naturaleza en la Alta Edad Media.

Este breve esbozo de la civilización en la Alta Edad Media, aunque bastante unilateral porque nos hemos referido sobre todo a su rela-

ción con las matemáticas, puede, sin embargo, proporcionar alguna idea sobre lo que era genuino de Europa y lo que Europa, construyendo sobre el exiguo legado de Roma, produjo bajo la guía de la Iglesia. Hasta 1100, el período medieval no produjo ninguna cultura grande en las esferas intelectuales. Las características de su situación intelectual eran las de falta de matización en las ideas, dogmatismo, misticismo, y confianza en las autoridades, que eran constantemente consultadas, analizadas y comentadas. Las inclinaciones místicas llevaban a la gente a elevar a realidades lo que eran vagas ideas, e incluso a aceptarlas como si fueran verdades religiosas. Lo poco que existía como ciencia teórica era estático. La Teología encerraba todo el conocimiento, y los Padres de la Iglesia desarrollaron sistemas de conocimiento universal. Pero no concibieron o buscaron principios diferentes de los contenidos en las doctrinas cristianas.

La civilización romana no fue productiva en matemáticas porque estaba demasiado preocupada por la obtención de resultados prácticos e inmediatamente aplicables. La civilización de la Europa medieval no lo fue tampoco, precisamente por la razón opuesta. No estaba preocupada en absoluto por el mundo físico. Los asuntos y problemas mundanos no tenían importancia. La Cristiandad puso todo el énfasis en la vida después de la muerte y en la preparación para esa vida.

Aparentemente, las matemáticas no pueden florecer ni en una civilización demasiado ligada a la tierra ni en una demasiado ligada al cielo. Tendremos oportunidad de ver que donde se han desarrollado con más éxito ha sido en una atmósfera intelectual libre que combine un interés por los problemas que presenta el mundo físico con un deseo de pensar sobre ideas sugeridas por esos problemas en una forma abstracta que no promete ningún resultado práctico o inmediato. La Naturaleza es la matriz de la que nacen las ideas. Esas ideas deben, entonces, estudiarse por ellas mismas. Así, paradójicamente, se obtiene una nueva visión de la naturaleza, una comprensión más rica, más amplia, más potente, lo que genera, a su vez, actividades matemáticas más profundas.

5. El primer renacimiento de las obras griegas

Hacia 1100 la civilización de Europa estaba, de alguna manera, estabilizada. Aunque la sociedad era fundamentalmente feudal, existían ya numerosos comerciantes independientes, se iniciaban la indus-

tria, la agricultura en gran escala, el manufacturado, la minería, los bancos, la ganadería, y gente libre o independiente cultivaba el arte y la artesanía. Se había establecido el comercio exterior, en particular con los árabes y el Oriente Próximo. Finalmente, tanto los príncipes como los dignatarios de la Iglesia y los comerciantes habían adquirido la riqueza necesaria para financiar la enseñanza y el cultivo de las artes. Aunque había ya una sociedad estable, hay pocos indicios de que, si hubieran seguido su propia forma de vida, los europeos hubieran abandonado alguna vez su concepto de la vida y el énfasis ya esbozado con respecto a ella, y se hubieran concentrado en un estudio serio de las matemáticas. La Europa Occidental era la sucesora de la Roma cristianizada, y ni Roma ni la Cristiandad se habían inclinado hacia las matemáticas. Pero alrededor de 1100 nuevas influencias comenzaron a afectar a la atmósfera intelectual. A través del comercio y de los viajes, los europeos se habían puesto en contacto con los árabes del área mediterránea y del Oriente Próximo y con los bizantinos del Imperio Romano de Oriente. Las Cruzadas (*c.* 1100-*c.* 1300), campañas militares para conquistar territorios, llevaron a los europeos a tierras árabes. Los cruzados eran hombres de acción más que de ilustración; quizá por ello la importancia de los contactos a través de las Cruzadas ha sido sobreestimada. De cualquier forma, los europeos comenzaron a conocer los trabajos griegos gracias a los árabes y a los griegos bizantinos.

La toma de conciencia del conocimiento griego creó una gran conmoción; los europeos se pusieron a buscar activamente copias de los trabajos griegos, de sus versiones árabes y textos escritos por árabes. Los príncipes y dignatarios de la Iglesia respaldaron a muchos eruditos en su búsqueda de esos tesoros. Estos eruditos fueron a los centros árabes de Africa, España, sur de Francia, Sicilia y Oriente Próximo para estudiar los trabajos existentes y llevarse consigo lo que pudieran comprar. Adelardo de Bath (*c.* 1090-*c.* 1150) fue a Siria, que estaba bajo control árabe, a Córdoba, disfrazado de estudiante mahometano, y al sur de Italia. Leonardo de Pisa aprendió aritmética en el norte de Africa. Las repúblicas del norte de Italia y el Papado enviaron comisiones y embajadores al Imperio Bizantino y a Sicilia, que era el origen de los famosos centros griegos y que, hasta el 878, había estado bajo el poder de Bizancio. En 1085 Toledo cayó en poder de los cristianos y se abrió allí para los eruditos europeos un centro importantísimo para el estudio de los trabajos árabes. Sicilia fue conquistada por los cristianos a los árabes en 1091, y los trabajos allí

existentes pudieron ser consultados libremente desde entonces. Investigaciones realizadas en Roma, que poseía obras griegos de los días del Imperio, permitieron sacar a la luz más manuscritos.

Al ir obteniendo esas obras, los europeos se propusieron, en forma creciente, ir traduciéndolos al latín. Las traducciones del griego del siglo XII no fueron, en general, buenas, porque el griego no se conocía bien. Eran *de verbo ad verbum;* pero eran mejores que las traducciones de las obras griegas que habían pasado a través del árabe, una lengua bastante poco parecida al griego. Por lo tanto, hasta bien entrado el siglo XVII, hubo una producción constante de nuevas y mejoradas traducciones.

Europa pudo, pues, conocer los trabajos de Euclides y Ptolomeo, la *Aritmética* y el *Algebra* de al-Kuo-ârizmî, la *Esférica* de Teodosio, muchos trabajos de Aristóteles y Herón, y un par de trabajos de Arquímedes, en particular su *Medida de una circunferencia.* (El resto de sus obras fue traducido al latín en 1544 por Hervagio de Basle.) Ni Apolonio ni Diofanto fueron traducidos durante los siglos XII y XIII. También fueron traducidos trabajos de filosofía, medicina, ciencia, teología y astrología. Como los árabes tenían casi todos los trabajos griegos, los europeos adquirieron una literatura inmensa. Admiraron tanto estas obras y les asombraron tanto las nuevas ideas que descubrieron en ellas, que se convirtieron en discípulos del pensamiento griego, y llegaron a valorar bastante más esos trabajos que sus propias creaciones.

6. El renacimiento del racionalismo y del interés por la naturaleza

Un enfoque racional de los fenómenos naturales y de su explicación en términos de causas naturales, en oposición a las explicaciones morales o de intenciones, comenzó a dar señales de vida casi inmediatamente después de que llegaran a Europa las primeras traducciones de los trabajos árabes y griegos. Un grupo de estudiosos en Chartres, Francia, constituido por Gilbert de la Porée (*c.* 1076-1154), Thierry de Chartres (m.*c.* 1155), y Bernard Sylvester (*c.* 1150), habían comenzado a buscar explicaciones racionales incluso de pasajes de la Biblia y hablaban, al menos, de la necesidad de utilizar las matemáticas en el estudio de la naturaleza. Sus doctrinas seguían el *Timeo* de Platón, pero eran más racionales que ese diálogo. Sin embargo, sus pronun-

ciamientos sobre los fenómenos físicos, aunque notables en el pensamiento medieval, no fueron ni significativos ni lo suficientemente influyentes como para dedicarles más atención aquí.

Con el influjo de los trabajos griegos se realzó la tendencia a las explicaciones racionales, al estudio del mundo físico y al interés en el disfrute del mundo real, a través de la comida, de la vida física y de los placeres de la naturaleza. Hubo quien incluso llegó a confrontar su propia razón contra la autoridad de la Iglesia. Así, Adelardo de Bath decía que no escucharía a quienes están «sujetos por las bridas...; por lo que si quieres comunicarte conmigo, intercambia razones».

Un hecho bastante sorprendente es que la introducción de algunos de los trabajos griegos retrasó la consciencia de Europa un par de siglos. Hacia 1200, los extensos trabajos de Aristóteles eran ya razonablemente conocidos. Los intelectuales europeos estaban complacidos e impresionados por su vasta acumulación de hechos, sus agudas distinciones, sus convincentes razonamientos y su organización lógica del conocimiento. El defecto en las doctrinas de Aristóteles era su aceptación de todo aquello que interesara a la mente, casi sin considerar su correspondencia con la experiencia. Ofrecía conceptos, teorías y explicaciones, como la doctrina de las sustancias básicas, la distinción entre los cuerpos terrestres y celestes (cap. 7, sec. 3), y el énfasis en la causa final, que tenían poca base en la realidad, o que no resultaron fructíferos. Como todas estas doctrinas fueron aceptadas sin discusión, las nuevas ideas, o no fueron consideradas, o no tuvieron suficiente audiencia, y se retrasó el posible progreso a que habrían conducido. También fue quizás un obstáculo el que Aristóteles asignara a las matemáticas un papel menor, ciertamente subordinado a la explicación física, la cual, para Aristóteles, era cualitativa.

El trabajo científico del período comprendido aproximadamente entre 1100 y 1450 fue realizado por los escolásticos, quienes se adhirieron a doctrinas basadas en la autoridad de los Padres Cristianos y de Aristóteles; su trabajo se resintió en consecuencia. Algunos de los escolásticos se rebelaron en contra del dogmatismo dominante y de la corrección absoluta de la ciencia de Aristóteles. Uno de los que sintieron la necesidad de obtener principios generales a partir de la experimentación, y de las deducciones en las que las matemáticas jugaran un papel y que pudieran entonces ser contrastadas con los hechos, fue el filósofo natural Robert Grosseteste (*c.* 1168-1253), obispo de Lincoln.

El disidente más elocuente contra la autoridad, que además tenía

ideas genuinas que ofrecer, fue Roger Bacon (1214-1294), el *Doctor Mirabilis*. Declaró: «Si tuviera poder sobre los trabajos de Aristóteles, los hubiera quemado todos; porque es sólo una pérdida de tiempo estudiarlos, y una causa de error, y una multiplicación de la ignorancia más allá de toda expresión.» Los enormes conocimientos de Bacon cubrían las ciencias de su tiempo y numerosas lenguas, incluyendo el árabe. Mucho antes de que resultaran ampliamente conocidos, él estaba informado de los últimos inventos y avances científicos: la pólvora, la acción de las lentes, los relojes mecánicos, la construcción del calendario y la formación del arco iris. Incluso comentó ideas sobre submarinos, aeroplanos y automóviles. Sus escritos sobre matemáticas, mecánica, óptica, visión, astronomía, geografía, cronología, química, perspectiva, música, medicina, gramática, lógica, metafísica, ética y teología fueron profundos.

Lo que es especialmente notable de Bacon es que comprendió cómo se obtiene un conocimiento fiable. Se preguntó por las causas que producen o impiden el avance de la ciencia, y especuló sobre la reforma de los métodos de investigación. Aunque recomendó el estudio de las Escrituras, atribuyó un gran valor a las matemáticas y a la experimentación, y previó importantes expectativas que podrían realizarse mediante la ciencia.

Las ideas matemáticas, afirma, son innatas en nosotros e idénticas a cosas tales como las que se encuentran en la naturaleza, porque la naturaleza está escrita en el lenguaje de la geometría. Por lo tanto, las matemáticas ofrecen la verdad. Son anteriores a las otras ciencias, porque permiten el conocimiento de la cantidad, la cual es aprehendida por la intuición. «Demuestra» en un capítulo de su *Opus Majus* que toda ciencia requiere matemáticas, y sus razonamientos muestran una justa apreciación del papel de las matemáticas en la ciencia. Aunque realza las matemáticas, también reconoce completamente el papel y la importancia de la experimentación como medio de descubrimiento y como comprobación de resultados obtenidos teóricamente, o de cualquier otra manera. «El razonamiento resuelve una pregunta; pero no nos hace sentirnos seguros, o asentir en la contemplación de la verdad, excepto cuando también se obtiene la verdad que lo es mediante la experiencia.»

La *Opus Majus* de Bacon trata bastante de la utilidad de las matemáticas en la geografía, la cronología, la música, la explicación del arco iris, el cálculo del calendario y la justificación de la fe. También trata sobre el papel de las matemáticas en la administración

del Estado, meteorología, hidrografía, astrología, perspectiva, óptica y visión.

Sin embargo, incluso Bacon era un producto de su tiempo. Creía en la magia y en la astrología, y mantenía que la teología era el objetivo de todo aprendizaje. Fue también una víctima de su tiempo: acabó en prisión, como muchos otros importantes intelectuales que habían comenzado a afirmar la prioridad e independencia de la razón humana y la importancia de la observación y de la experimentación. Su influencia en su tiempo no fue grande.

Guillermo de Ockham (*c.* 1300-1349) continuó los fuertes ataques a Aristóteles, criticando los puntos de vista de éste sobre las causas. La causa final, decía, es una pura metáfora. Todas las causas son inmediatas, y la causa total es el agregado de todas las antecedentes que bastan para que se realice un suceso. Este conocimiento de las conexiones tiene una validez universal por la uniformidad de la naturaleza. La función primaria de la ciencia es establecer sucesiones de observaciones. Por lo que se refiere a la sustancia, Ockham decía que conocemos sólo propiedades, no una forma sustancial fundamental.

También atacó la física y la metafísica contemporáneas, diciendo que el conocimiento obtenido de la experiencia es real, mientras que las construcciones racionales no lo son; éstas están inventadas sólo para explicar los hechos observados. Su famoso principio es «la navaja de Ockham» —ya establecido por Grosseteste y Duns Scoto (1266-1308)—: es fútil trabajar con más entidades cuando basten menos. Separó la teología de la filosofía natural (ciencia), basado en que la teología obtiene conocimientos a partir de revelaciones, mientras que la filosofía natural debe obtenerlos mediante la experiencia.

Estos disidentes no sugirieron nuevas ideas científicas. Sin embargo, presionaron para obtener una mayor libertad de especulación, pensamiento e investigación, y propiciaron la experiencia como la fuente del conocimiento científico.

7. El progreso específico en matemáticas

A pesar de los rígidos límites del pensamiento en el período comprendido aproximadamente entre 1100 y 1450, tuvo lugar alguna actividad matemática. Los principales centros en los que se desarrolló fueron las universidades de Oxford, París, Viena (fundada en 1365) y

Erfurt (fundada en 1392). El trabajo inicial fue una respuesta directa a la literatura griega y árabe.

El primer europeo que merece mencionarse es Leonardo de Pisa (*c.* 1170-1250), también llamado Fibonacci. Fue educado en Africa, viajó extensamente por Europa y Asia Menor, y fue famoso por su soberana posesión de todo el conocimiento matemático de su generación y de las precedentes. Residió en Pisa, y era bien conocido de Federico II de Sicilia y de los filósofos de la corte, a quienes están dedicados muchos de sus trabajos.

En 1202, Leonardo escribió su *Liber Abaci*, muy usado y que hizo época, y que consistía en una traducción libre de materiales árabes y griegos al latín. La notación árabe para los números, y los métodos de cálculo hindúes, ya eran algo conocidos en Europa, pero sólo en los monasterios. La gente utilizaba en general los numerales romanos, y evitaban el cero porque no lo entendían. El libro de Leonardo ejerció una gran influencia y cambió el panorama; enseñó el método hindú de cálculo con enteros y fracciones, raíces cuadradas y raíces cúbicas. Estos métodos fueron mejorados después por comerciantes florentinos.

Tanto en el *Liber Abaci* como en un trabajo posterior, el *Liber Quadratorum* (1225), Leonardo se ocupó del álgebra. Siguió a los árabes en usar palabras en lugar de símbolos y basar el álgebra en métodos aritméticos. Expuso la solución de ecuaciones determinadas e indeterminadas de primero y segundo grado, así como de algunas ecuaciones cúbicas. Al igual que Khayyam, creía que las ecuaciones cúbicas no podían ser resueltas algebraicamente.

Referente a la geometría, Leonardo, en su *Practica Geometriae* (1220), reprodujo buena parte de los *Elementos* de Euclides y de la trigonometría griega. Sus enseñanzas de agrimensura mediante métodos trigonométricos en lugar de los métodos geométricos romanos representó un ligero avance.

La característica nueva más significativa del trabajo de Leonardo es la observación de que la clasificación de Euclides de los irracionales en el libro X de los *Elementos* no incluía todos los irracionales. Leonardo mostró que las raíces de $x^3 + 2x^2 + 10x = 20$ no pueden construirse con regla y compás. Esta fue la primera indicación de que el sistema de números contenía más de los que permitía el criterio griego de existencia basado en la construcción mencionada. Leonardo introdujo también la noción de lo que todavía se llama sucesiones de Fibonacci, en las que cada término es la suma de los dos precedentes.

Además de la observación de Leonardo sobre los irracionales, había algunas ideas germinales en el trabajo de Nicolás Oresme (*c.* 1323-1382), obispo de Lisieux y profesor en el Colegio Parisino de Navarra. En su trabajo no publicado *Algorismus Proportionum* (*c.* 1360), presentó tanto una notación como algunos cálculos para exponentes fraccionarios. Su reflexión fue que, puesto que (en nuestra notación) $4^3 = 64$ y que $(4^3)^{1/2} = 8$, entonces $4^{3/2} = 8$. La noción de exponentes fraccionarios reapareció en los trabajos de varios escritores del siglo XVI, pero no fue utilizada ampliamente hasta el XVII.

Otra contribución del trabajo de Oresme fue el estudio del cambio. Recordemos que Aristóteles distinguía nítidamente entre cualidad y cantidad. La intensidad del calor era una cualidad. Para cambiar la intensidad, una sustancia, una especie de calor, debe perderse y otra añadirse. Oresme afirmaba que no había tipos diferentes de calor sino más o menos cantidad del mismo tipo. Varios escolásticos del siglo XIV en Oxford y París, incluyendo a Oresme, comenzaron a pensar sobre el cambio y la velocidad del cambio cuantitativamente. Estos autores estudiaron el movimiento uniforme (movimiento con velocidad constante), movimiento diforme (movimiento con velocidad variable) y movimiento uniformemente diforme (movimiento con aceleración constante).

Esta línea de pensamiento culminó en ese período con la doctrina de Oresme de la latitud de las formas. Sobre este tema escribió *De Uniformitate et Difformitate Intensionum* (*c.* 1350) y *Tractatus Latitudinibus Formarum*. Para estudiar el cambio y la velocidad del cambio, Oresme siguió la tradición griega afirmando que las cantidades medibles distintas de números podían representarse mediante puntos, líneas y superficies. Por ello, para representar el cambio de la velocidad con el tiempo, representa el tiempo a lo largo de una línea horizontal, que llama la longitud, y las velocidades en distintos instantes de tiempo mediante líneas verticales, que llamó latitudes. Para representar una velocidad que decrece uniformemente desde el valor *OA* (fig. 10.1) en *O* a cero en *B*, dibuja una figura triangular. También apunta que el rectángulo *OBDC*, determinado por *E*, el punto medio de *AB*, tiene la misma área que el triángulo *OAB* y representa un movimiento uniforme a lo largo del mismo intervalo de tiempo. Oresme asociaba el cambio físico con toda la figura geométrica. El área completa representaba la variación en cuestión; no había referencia a valores numéricos.

Se dice a menudo que Oresme contribuyó a la formación del

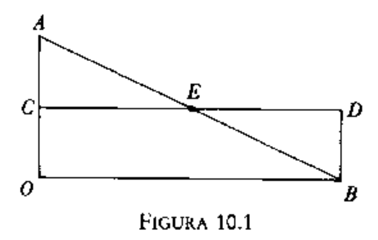


FIGURA 10.1

concepto de función, a la representación funcional de las leyes físicas, y a la clasificación de las funciones. Se le ha acreditado también la creación de la geometría de coordenadas y la representación gráfica de funciones. En realidad la latitud de las formas es una idea poco clara, como mucho un tipo de gráfico. Aunque la representación de Oresme de la intensidad bajo el nombre de *latitudines formarum* fue una técnica importante en el objetivo escolástico de estudiar el cambio físico, y fue aplicada en algunos intentos de revisar la teoría de Aristóteles sobre el movimiento, su influencia sobre el pensamiento posterior fue pequeña. Galileo utilizó esta figura, pero con bastante más claridad y acierto. Como Descartes evitó cuidadosamente toda referencia a sus predecesores, no sabemos si fue influido por las ideas de Oresme.

8. El progreso en las ciencias físicas

Ya que el progreso de las matemáticas depende vitalmente de la renovación del interés por la ciencia, anotaremos brevemente los esfuerzos realizados en este área por el hombre medieval.

En mecánica, incorporaron los muy aceptables trabajos griegos sobre la palanca y los centros de gravedad, y los trabajos de Arquímedes sobre hidrostática. Hicieron poco más que comprender la teoría de la palanca, aunque Jordanus Nemorarius (m. 1237) realizó algunas ligeras aportaciones. La que recibió más atención fue la teoría del movimiento.

Como la ciencia de Aristóteles había adquirido cierto ascendiente, fue su teoría la que constituyó el punto de partida para las investigaciones sobre el movimiento. Como apuntamos en el capítulo 7, había varias dificultades aparentes en la teoría de Aristóteles. Los primeros científicos medievales intentaron resolverlas dentro del marco de

referencia aristotélico. Por ejemplo, para interpretar la aceleración de la caída de los cuerpos, varios estudiosos del siglo XIII interpretaban las vagas nociones de Aristóteles sobre la gravedad queriendo significar que el peso de un cuerpo aumenta cuando se aproxima al centro de la Tierra. Entonces, puesto que la fuerza aumenta, también lo hace la velocidad. Otros cuestionaron la ley fundamental de Aristóteles de que la velocidad es la fuerza dividida por la resistencia.

A la escuela de Chartres sucedió en el siglo XIV una escuela de París, cuyas cabezas fueron Oresme y Jean Buridan (*c.* 1300-*c.* 1360). Aquí, en la universidad, las concepciones aristotélicas habían sido dominantes. Para explicar el movimiento continuo de los objetos lanzados por una fuerza, Buridan desarrolló una nueva teoría, la teoría del ímpetu. Siguiendo al erudito cristiano del siglo VI Filopón, Buridan decía que la potencia motriz aplicada a una flecha o a un proyectil estaba aplicada al objeto mismo y no al aire. Este ímpetu, más que la acción propulsora del aire, mantendría la velocidad uniforme del objeto indefinidamente, en ausencia de fuerzas externas. En el caso de los cuerpos que caen, el ímpetu aumentaría gradualmente por la actuación de la gravedad natural, que añadiría sucesivos incrementos de ímpetu al que el objeto ya tenía. En cuerpos ascendentes —por ejemplo, proyectiles— el ímpetu proporcionado al objeto disminuiría gradualmente por la resistencia del aire y la gravedad natural. Impetu era lo que Dios había dado a las esferas celestes, que no necesitaban agentes celestiales para mantenerlas en movimiento. Buridan definió el ímpetu como la cantidad de materia multiplicada por la velocidad; en consecuencia, en términos modernos, es el momento.

Esta nueva teoría era notable por varias razones. Aplicándola a los movimientos en el cielo y la tierra, Buridan pudo unirlos en una teoría. Además, la teoría implicaba que, contrariamente a la ley de Aristóteles, una fuerza podía alterar el movimiento y no sólo mantenerlo. Tercero, el concepto de ímpetu en sí mismo era un gran avance; transfería la potencia motriz del medio al objeto móvil y también hacía posible la consideración de un vacío. Buridan es uno de los fundadores de la dinámica moderna. Su teoría fue ampliamente aceptada en su propio siglo y durante los dos siglos siguientes.

Quizá el movimiento de los proyectiles recibió esta atención porque el armamento mejorado del siglo XIII, como catapultas, ballestas y arcos, podía lanzar proyectiles a lo largo de trayectorias más largas y más arqueadas; un siglo más tarde se utilizaban los

cañones. Aristóteles había dicho que un cuerpo sólo puede moverse bajo la acción de un tipo de fuerza en cada momento; si estuvieran actuando dos, una destruiría a la otra. Por lo tanto, un cuerpo arrojado hacia arriba y hacia afuera se movería a lo largo de una línea recta hasta que el movimiento «violento» se consumiera y entonces el cuerpo caería directamente a la tierra según un movimiento natural. De las diversas revisiones de esta teoría, la idea de Jordano Nemorario resultó ser la más útil; mostró que la fuerza bajo la que se mueve un cuerpo arrojado en línea recta, puede descomponerse en cualquier instante en dos componentes, la gravedad natural actuando hacia abajo y una fuerza de proyección «violenta» horizontal. Esta idea fue adoptada por da Vinci, Stevin, Galileo y Descartes.

La escuela parisina de Buridan y Oresme consideró no sólo el movimiento uniforme sino también el diforme y el uniformemente diforme, y demostró para su propia satisfacción que la velocidad efectiva en el movimiento uniformemente diforme era la media de las velocidades inicial y final. Quizá lo más significativo acerca de los esfuerzos realizados en el desarrollo de la mecánica en los siglos XIII y XIV fue el intento de introducir consideraciones cuantitativas y reemplazar razonamientos cualitativos por cuantitativos.

El principal interés de los científicos medievales se centraba en el campo de la óptica. Una razón para ello es que los griegos habían establecido, en lo que hoy llamamos óptica geométrica, unos fundamentos más firmes que en cualquier otro campo de la física, y a finales del período medieval sus numerosos trabajos en óptica eran conocidos en Europa. Además, los árabes habían obtenido algunos avances adicionales. Hacia 1200, algunas de las leyes básicas sobre la luz eran bien conocidas, como la de la propagación en línea recta de la luz en un medio uniforme, la ley de la reflexión y la incorrecta ley de Ptolomeo sobre la refracción (creía que el ángulo de refracción era proporcional al ángulo de incidencia). También habían pasado a Europa de los griegos y los árabes conocimientos como el de los espejos esféricos y parabólicos, la aberración esférica, la cámara oscura, los usos de las lentes, el funcionamiento del ojo, la refracción atmosférica y la posibilidad de ampliación.

Los científicos Grosseteste, Roger Bacon, Vitello (siglo XIII), John Peckham (m. 1292) y Teodorico de Freiberg (m.*c.* 1311) aportaron algunos avances en el estudio de la luz. Habiendo estudiado la refracción de la luz a través de lentes, determinaron las longitudes focales de algunas lentes, estudiaron combinaciones de lentes, sugirie-

ron la ampliación de las imágenes mediante combinaciones de lentes y aportaron mejoras a la teoría del arco iris. Los espejos de vidrio fueron perfeccionados en el siglo XIII; las gafas se remontan a 1299. Vitello observó la dispersión de la luz bajo refracción; esto es, produjo colores haciendo pasar luz blanca a través de un cristal hexagonal. También hizo pasar la luz a través de un cuenco de agua para estudiar el arco iris, pues había notado previamente que cuando la luz del sol pasaba a través de un cuenco de agua los colores del arco iris aparecían en la luz emergente. La óptica continuó siendo una ciencia importante; encontraremos a Kepler, Galileo, Descartes, Fermat, Huygens y Newton trabajando en ella.

9. Sumario

En la ciencia, como en otros campos, la Edad Media se concentró en los trabajos consagrados y contrastados por el tiempo. Las escuelas produjeron adecuadas selecciones de antiguos manuscritos, resúmenes y comentarios. El espíritu de los tiempos obligaba a las mentes a seguir caminos rígidos, prescritos, sólidos. La búsqueda de una filosofía universal que cubriera todos los fenómenos del hombre, de la naturaleza y de Dios es característica del último período medieval. Pero las contribuciones adolecieron de una falta de matización en las ideas, de misticismo, de dogmatismo y de un espíritu de comentario dirigido al análisis de lo que se consideraba la autoridad.

Sin embargo, a medida que el mundo cambiaba, fue creciendo cada vez con más fuerza la consciencia de la discrepancia y el conflicto entre las creencias y los hechos manifiestos, y se hizo más clara la necesidad de una revisión del aprendizaje y de las creencias. Antes de que Galileo demostrara el valor de la experiencia, antes de que Descartes enseñara a la gente a mirarse a sí mismos y antes de que Pascal formulara la idea de progreso, fueron los pensadores no convencionales, en su mayor parte los escolásticos disidentes, los que intentaron avanzar a lo largo de nuevas líneas, desafiando perspectivas establecidas y confiando más fuertemente en la observación de la naturaleza de lo que lo habían hecho los griegos.

La experimentación, motivada en parte por la búsqueda de poderes mágicos, y el uso de la inducción para obtener principios generales o leyes científicas, comenzaron a ser fuentes importantes del conocimiento, a pesar del hecho de que el principal método científico de la

Baja Edad Media era la explicación racional, presentada en una demostración formal o geométrica basada en principios a priori.

El valor de las matemáticas en la investigación de la naturaleza también recibió algún reconocimiento. Aunque, en general, los científicos medievales siguieron a Aristóteles en su búsqueda de explicaciones materiales o físicas, éstas eran difíciles de obtener y no fueron demasiado útiles; encontraron paulatinamente que era más fácil relacionar observaciones y resultados experimentales matemáticamente y entonces comprobar la ley matemática relevante. Por ello, en astronomía, los científicos interesados en la teoría propiamente dicha, en la navegación y en el calendario no utilizaron la modificación de Aristóteles de la teoría de Eudoxo, sino la teoría de Ptolomeo. Como consecuencia, las matemáticas comenzaron a jugar un papel mayor que el que Aristóteles les había asignado.

A pesar de las nuevas tendencias y actividades, es dudoso que la Europa medieval, si hubiera proseguido un camino sin cambios, hubiera desarrollado nunca una auténtica ciencia y las matemáticas. La investigación libre no estaba permitida. Las pocas universidades ya existentes hacia 1400 estaban controladas por la Iglesia, y los profesores no eran libres para enseñar lo que ellos estimaban correcto. Si no se condenaron doctrinas en el período medieval fue sólo porque no se estableció ninguna importante. Cualquier disensión verdadera del pensamiento cristiano que apareciera en cualquier esfera era suprimida sumarísima y cruelmente, y con una malicia sin igual en la historia, sobre todo a través de la Inquisición, iniciada por el Papa Inocencio III en el siglo XIII.

Otros factores, relativamente menos importantes, retrasaron también los cambios en Europa. El renacido conocimiento griego alcanzó sólo a los pocos estudiosos que disponían del ocio y de la oportunidad para estudiarlo. Los manuscritos eran caros; muchos de los que querían comprarlos no podían. Además Europa, en el período comprendido entre 1100 y 1500, estaba dividida en numerosos ducados, principados, ciudades-estado más o menos democráticas u oligárquicas y los Estados Pontificios, todos ellos independientes. Las guerras entre todas esas unidades políticas eran continuas y absorbían las energías de la gente. Las Cruzadas, que empezaron alrededor de 1100, desperdiciaron un enorme número de vidas. La Peste Negra, en la segunda mitad del siglo XIV, se llevó aproximadamente la tercera parte de la población de Europa, e hizo retroceder la civilización en su conjunto. Afortunadamente, fuerzas de tendencia revolucionaria co-

menzaron a ejercer su influencia en las escenas social, política e intelectual europeas.

Bibliografía

Ball, W. W. R.: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, caps. 8, 10, 11.
Boyer, Carl B.: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.
Cajori, Florian: *A History of Mathematics*, Macmillan, 1919, pp. 113-129.
Claggett, Marshall: *The Science of Mechanics in the Middle Ages*, University of Wisconsin Press, 1959.
—: *Nicole Oresme and the Geometry of Qualities and Motions*, University of Wisconsin Press, 1968.
Crombie, A. C.: *De Agustín a Galileo*, Madrid, Alianza, 1979.
—: *Robert Grosseteste and the Origins of Experimental Science*, Oxford University Press, 1953.
Easton, Stewart: *Roger Bacon and His Search for a Universal Science*, Columbia University Press, 1952.
Hofmann, J. E.: *The History of Mathematics*, Philosophical Library, 1957, caps. 3-4.
Smith, David Eugene: *History of Mathematics*, Dover (reimpresión), 1958, vol. I, pp. 177-265.

Capítulo 11
EL RENACIMIENTO

> Me parece que si alguien quiere avanzar en matemáticas debe estudiar a los maestros y no a los discípulos.
>
> N. H. ABEL

1. Influencias revolucionarias en Europa

En el período comprendido aproximadamente entre 1400 y 1600, que adoptaremos como el período del Renacimiento (aunque este término se utiliza para describir diferentes períodos cronológicos por distintos autores), Europa fue sacudida profundamente por una cantidad de acontecimientos que acabaron por alterar drásticamente las perspectivas intelectuales, y agitaron la actividad matemática con magnitud y profundidad sin precedentes.

Las influencias revolucionarias se difundieron ampliamente. Los cambios políticos que sucedieron a las casi incesantes guerras afectaron a todas las ciudades y estados de Europa. Italia, la madre del Renacimiento, es un ejemplo importante. Aunque la historia de los estados italianos en los siglos XV y XVI está plagada de intrigas constantes, asesinatos en masa, y destrucción provocada por las guerras, la naturaleza muy fluida de las condiciones políticas y el establecimiento de algunos gobiernos democráticos favorecieron el desarrollo del individuo. Las guerras contra el papado, una importante potencia política y militar en esa época, no sólo liberaron al pueblo de la dominación de la Iglesia, sino que estimularon también la oposición intelectual.

Italia acumuló una gran riqueza durante el último período de la Edad Media. Esto se debió principalmente a su situación geográfica. Sus puertos de mar eran los más favorablemente situados para importar artículos de Africa y Asia y enviarlos al resto de Europa. Grandes casas de banca hicieron de Italia el centro financiero. Esta riqueza era esencial para propiciar el aprendizaje. Y fue en Italia, el más atormentado de los países, hirviendo en tumultos, donde primero se concibieron y expresaron los modos de pensar que iban a moldear la civilización occidental.

Durante el siglo XV llegaron a Europa, en enormes cantidades, las obras griegas. Durante la primera parte del siglo se hicieron más fuertes las conexiones entre Roma y el Imperio Bizantino, que poseía la mayor colección de documentos griegos, pero que había estado muy aislado. Los bizantinos, en guerra con los turcos, buscaron el apoyo de los estados italianos. Con estas mejores relaciones, profesores griegos fueron a Italia, e italianos al Imperio Bizantino para aprender griego. Cuando los turcos conquistaron Constantinopla en 1453, los estudiosos griegos marcharon a Italia, llevando más manuscritos con ellos. Por tanto, no sólo hubo más manuscritos griegos disponibles en Europa, sino que muchos de los recientemente adquiridos eran mucho mejores que los adquiridos previamente en los siglos XII y XIII. Las traducciones ulteriores, realizadas al latín directamente del griego, eran más fiables que las realizadas del árabe.

La invención de Johann Gutenberg, hacia 1450, de la imprenta con tipos movibles, aceleró la difusión del conocimiento. El papel de lino y algodón, que los europeos heredaron de los chinos a través de los árabes, sustituyó al pergamino y al papiro a partir del siglo XII. Desde 1474, los trabajos matemáticos, astronómicos y astrológicos aparecieron en forma impresa. Por ejemplo, la primera edición impresa de los *Elementos* de Euclides en una traducción latina realizada por Johannes Campanus (siglo XIII), apareció en Venecia en 1482. A lo largo del siglo siguiente, aparecieron en ediciones impresas los primeros cuatro libros de las *Secciones Cónicas* de Apolonio, los trabajos de Pappus, la *Arithmetica* de Diofanto, así como otros tratados.

La utilización de la brújula y de la pólvora tuvo efectos significativos. La brújula hizo posible la navegación más allá del alcance de la vista desde tierra. La pólvora, introducida en el siglo XIII en Europa, cambió los métodos de guerra y el diseño de las fortificaciones, e hizo que resultara importante el estudio del movimiento de proyectiles.

Se inició una nueva era económica, debido a un crecimiento

enorme en el manufacturado, minería, agricultura en gran escala y una gran variedad de oficios. Cada una de estas empresas encontró problemas técnicos que fueron tratados más vigorosamente que en cualquier otra civilización previa. En contraste con las sociedades esclavas de Egipto, Grecia y Roma, y la servidumbre de la gleba del feudalismo, la nueva sociedad poseía una clase en expansión de obreros y artesanos libres. Trabajadores independientes y patronos con asalariados tenían incentivos para pensar e inventar mecanismos que ahorraran trabajo. La competitividad de una economía capitalista también estimuló estudios directos de fenómenos físicos y conexiones causales para mejorar materiales y métodos de producción. Como la Iglesia había proporcionado explicaciones de muchos de esos fenómenos, surgieron los conflictos. Podríamos estar seguros de que siempre que la explicación física se revelaba más útil que la teológica, ésta era ignorada.

La clase de los comerciantes contribuyó a instaurar un nuevo orden en Europa promoviendo las exploraciones geográficas de los siglos XV y XVI. Inducidas por la necesidad de disponer de mejores rutas para el comercio y de proveedores de mercancías, las exploraciones trajeron a Europa mucho conocimiento de tierras extrañas, plantas, animales, climas, formas de vida, creencias y costumbres.

Las dudas sobre la solidez de la ciencia y de la cosmología de la Iglesia suscitadas por la observación directa o por la información filtrada a Europa por exploradores y mercaderes, las objeciones a la supresión, por parte de la Iglesia, de la experimentación y la reflexión sobre los problemas creados por el nuevo orden social, la degeneración de algunos dignatarios de la Iglesia, las prácticas corruptas de la Iglesia, como la cuestión de las indulgencias y, finalmente, las serias diferencias doctrinales, culminaron en la Reforma protestante. Los reformistas fueron apoyados por comerciantes y príncipes, ansiosos por quebrar el poder de la Iglesia.

La Reforma como tal no liberalizó el pensamiento ni liberó las mentes. Los dirigentes protestantes sólo querían establecer su propio dogmatismo. Sin embargo, al suscitar cuestiones referentes a la naturaleza de los sacramentos, a la autoridad del gobierno de la Iglesia, y al significado de algunos pasajes de las Escrituras, Lutero, Calvino y Zuinglio involuntariamente estimularon a mucha gente para pensar y hacer lo que, de otra manera, ni siquiera hubieran intentado. Se estimuló el pensamiento y se provocó la discusión. Además, para ganar adeptos, los protestantes afirmaron que el juicio individual, en

vez de la autoridad papal, era la base de las creencias. Se consolidaron, pues, modificaciones en estas creencias. Muchos, puestos a elegir entre lo que decían los católicos o los protestantes, se desentendían de ambos y se volvían hacia las dos fes de la naturaleza, la observación y la experimentación, como fuentes del conocimiento.

2. La nueva perspectiva intelectual

La Iglesia se basaba en la autoridad, reverenciaba a Aristóteles, y calificaba la duda de criminal. También desaprobaba las satisfacciones materiales y propugnaba la salvación del alma para una vida futura. Estos principios contrastaban profundamente con los valores que los europeos habían aprendido de los griegos, aunque no se hubieran establecido explícitamente en los trabajos de Aristóteles: el estudio de la naturaleza, el disfrute del mundo físico, el perfeccionamiento de la mente y del cuerpo, la libertad de investigación y de expresión y la confianza en la razón humana. Molestos por la autoridad de la Iglesia, por las restricciones en la vida material y por la confianza en las Escrituras como la única fuente de todo conocimiento, los intelectuales se aferraron ansiosamente a los nuevos valores. En lugar de disputas sin fin y de discusiones sobre el significado de los pasajes bíblicos para la determinación de la verdad, los hombres volvieron la vista a la naturaleza misma.

Casi como corolario del Renacimiento y de los valores griegos, sobrevino el renacimiento del interés por las matemáticas. Especialmente de los trabajos de Platón, que se conocieron en el siglo XV, los europeos aprendieron que la naturaleza está diseñada matemáticamente, y que este diseño es armonioso, estéticamente agradable y la verdad última. La naturaleza es racional, simple y ordenada, y actúa según leyes inmutables. Las obras platónicas y pitagóricas realzaban el número como la esencia de la realidad, doctrina que ya había recibido alguna atención de los escolásticos desviacionistas de los siglos XIII y XIV. El renacimiento del platonismo clarificó y cristalizó las ideas y métodos con los que estos hombres habían estado luchando. El énfasis pitagórico-platónico sobre las relaciones cuantitativas como la esencia de la realidad se hizo dominante en forma gradual. Copérnico, Kepler, Galileo, Descartes, Huygens y Newton eran, en este aspecto, pitagóricos y, mediante su trabajo, establecieron el

principio de que el objetivo de la actividad científica debe ser la obtención de leyes matemáticas cuantitativas.

Las matemáticas atraían a los intelectuales del Renacimiento todavía por otra razón. El Renacimiento fue un período en el que la civilización y cultura medievales se iban desacreditando a medida que las nuevas influencias, información y movimientos revolucionarios recorrían Europa. Estos hombres buscaban bases nuevas y más firmes para edificar el conocimiento, y las matemáticas les ofrecían esos cimientos. Las matemáticas permanecieron como el único cuerpo de verdades aceptado en medio de sistemas filosóficos que se desmoronaban, de creencias teológicas discutidas y de valores éticos cambiantes. El conocimiento matemático era cierto, y ofrecía una base segura en un pantano; la búsqueda de la verdad fue reconducida hacia él.

Los matemáticos y los científicos recibieron alguna inspiración de los prejuicios teológicos de la Edad Media, que habían inculcado la visión de que todos los fenómenos de la naturaleza están no sólo interconectados, sino que se producen de acuerdo con un plan global: todas las acciones de la naturaleza siguen el plan establecido por una única causa primera. ¿Cómo, pues, se reconciliaba la visión teológica del universo de Dios con la búsqueda de las leyes matemáticas de la naturaleza? La respuesta fue una nueva doctrina, según la cual Dios había diseñado el universo matemáticamente. En otras palabras, atribuyendo a Dios el carácter de matemático supremo, se hacía posible dotar de un sentido religioso a la búsqueda de las leyes matemáticas de la naturaleza.

Esta doctrina inspiró el trabajo de los matemáticos de los siglos XVI y XVII, e incluso de algunos del siglo XVIII. La investigación de las leyes matemáticas de la naturaleza era un acto de devoción; era el estudio de los caminos y naturaleza de Dios y de su plan del universo. El científico del Renacimiento era un teólogo, cuyo tema era la naturaleza en vez de la Biblia. Copérnico, Brahe, Galileo, Pascal, Descartes, Newton y Leibniz hablan repetidamente de la armonía que Dios imprimió al universo mediante su diseño *matemático*. El conocimiento matemático, como es en sí mismo la verdad acerca del universo, es tan sacrosanto como cualquier línea de las Escrituras, e incluso superior, porque es un conocimiento claro e indiscutible. Galileo dijo: «No se descubre menos admirablemente Dios a nosotros en las acciones de la naturaleza que en las sagradas afirmaciones de las Escrituras.» A lo que Leibniz añadió: «*Cum Deus calculat, fit mundus*» (Como Dios calcula, así está hecho el mundo). Estos hombres

buscaban relaciones matemáticas que revelarían la magnificencia y gloria del trabajo de Dios. El hombre no podría esperar comprender el plan divino tan claramente como el mismo Dios lo entendía pero, con humildad y modestia, podría intentar al menos aproximarse a la mente de Dios.

Los científicos persistían en la búsqueda de leyes matemáticas que estuvieran en la base de los fenómenos naturales, porque estaban convencidos de que Dios había incorporado esas leyes en su construcción del universo. Cada descubrimiento de una ley de la naturaleza era aclamado más como prueba de la brillantez de Dios que como prueba de la del investigador. Kepler, en particular, escribió alabanzas a Dios con ocasión de cada descubrimiento. Las creencias y actitudes de los matemáticos y científicos ejemplifican el fenómeno cultural más amplio que recorrió la Europa renacentista. Las obras griegas incidieron en un mundo cristiano profundamente devoto, y los dirigentes intelectuales nacidos en uno y atraídos por el otro fusionaron las doctrinas de ambos.

3. La difusión de la instrucción

Por diversas razones, la difusión de los nuevos valores tuvo lugar lentamente. Las obras griegas podían encontrarse, al principio, sólo en las cortes de príncipes seculares o eclesiásticos, y no eran accesibles a la gente en general. La imprenta contribuyó enormemente a hacer los libros disponibles en general, pero el efecto fue gradual porque incluso los libros impresos eran caros. El problema de difundir el conocimiento era también complicado por dos factores adicionales. En primer lugar, muchos a quienes hubiera gustado utilizar las matemáticas y la ciencia en la industria, artesanía, navegación, arquitectura y otros fines, no eran cultos; la escolarización no era en absoluto común. El segundo factor fue la lengua. A las personas instruidas —eruditos, profesores y teólogos— les resultaba familiar el latín y, hasta cierto punto, el griego. Pero los artistas, artesanos e ingenieros conocían sólo las lenguas vernáculas —francés, alemán, y las diversas lenguas italianas— y no se beneficiaron de las traducciones latinas de las obras griegas.

A partir del siglo XVI, muchos clásicos griegos fueron traducidos a las lenguas populares. Los mismos matemáticos participaron en esta actividad. Por ejemplo, Tartaglia tradujo los *Elementos* de Euclides

del latín al italiano en 1543. El movimiento de traducción continuó hasta bien entrado el siglo XVII, pero se realizó lentamente porque muchos estudiosos eran hostiles y despreciaban al pueblo llano. Preferían el latín, porque creían que el peso de la tradición asociado a él conferiría autoridad a sus pronunciamientos. Para oponerse a estos estudiosos, llegar al público, y obtener apoyo para sus ideas, Galileo escribió deliberadamente en italiano. Descartes escribió en francés por la misma razón; esperaba que quienes utilizaran su razonamiento natural podrían ser mejores jueces de sus enseñanzas que los que respetaban servilmente los escritos antiguos.

Otra medida que se tomó para educar al público, adoptada en algunas ciudades de Italia, fue el establecimiento de bibliotecas. La familia Medici financió bibliotecas en Florencia y algunos papas hicieron lo mismo en Roma. En parte para difundir la educación y en parte para proporcionar un lugar de encuentro para la comunicación entre los estudiosos, algunos gobernantes liberales fundaron las academias. La más notable entre éstas fue la Academia Florentina, fundada por Cosimo I de Medici (1519-1574) en 1563, que se convirtió en un centro de estudios matemáticos, y la Academia Romana dei Lincei (de los linces) fundada en 1603. Miembros de estas academias traducían trabajos latinos a lenguas populares, daban conferencias al público y profundizaban en su propio conocimiento comunicándose con los demás. Estas academias fueron las precursoras de las más famosas, fundadas más tarde en Inglaterra, Francia, Italia y Alemania, y que fueron tan enormemente útiles en la difusión del conocimiento.

Lamentablemente, las universidades de los siglos XV y XVI casi no jugaron ningún papel en estos avances. La teología gobernaba las universidades y su estudio era el propósito del aprendizaje. El conocimiento era considerado completo y cerrado. Por lo tanto, la experimentación era innecesaria, e ignorados los nuevos descubrimientos de los de fuera. Lo más que hicieron los conservadores profesores de la universidad fue aferrarse al conocimiento medieval tal como había sido formulado por los escolásticos desde el siglo XIII. Las universidades enseñaban aritmética, geometría, astronomía y música, pero la astronomía estaba basada en la de Ptolomeo y no era observacional. La filosofía natural significaba el estudio de la *Física* de Aristóteles.

4. La actividad humanística en las matemáticas

Mientras que los escolásticos se adhirieron firmemente a las últimas doctrinas medievales, un nuevo grupo de humanistas se dedicó a reunir, organizar y estudiar críticamente el conocimiento griego y romano. Estos hombres realizaron estudios asiduos y, con una sagacidad cuestionable en su conjunto, se esforzaron por limpiar los textos de errores y por recuperar material perdido. Aceptaron servilmente, repitieron e interpretaron sin parar lo que habían encontrado en manuscritos antiguos y medievales, emprendiendo incluso estudios filosóficos para precisar los significados. Escribieron también muchos libros en los que rehicieron muchos trabajos antiguos en una reinterpretación escolástica. Aunque esta actividad pudo haber despertado algún interés por aprender, propició, sin embargo, la decepción de que el aprendizaje consistía en profundizar y confirmar el cuerpo de conocimiento aceptado.

Típico entre los humanistas del siglo XVI fue Gerónimo Cardano (Gerolamo Cardano), que nació en Pavía en 1501. Su carrera como pícaro y estudioso fue una de las más fascinantes entre las fantásticas trayectorias de los hombres del Renacimiento. Proporciona un relato de su vida en su obra *De vita propria* (De la propia vida), escrita cuando era ya anciano, en la que se ensalza y humilla a sí mismo. Dice que sus padres le dotaron sólo de miseria y desprecio; pasó una infancia miserable y fue tan pobre durante los primeros cuarenta años de su vida que dejó de considerarse pobre a sí mismo porque, según decía, no le quedaba nada que perder. Era de gran temperamento, dedicado a los placeres eróticos, pendenciero, engreído, sin sentido del humor, incapaz de sentir remordimiento e intencionadamente cruel en su manera de hablar. Aunque no era un apasionado del juego, jugó a los dados todos los días durante veinticinco años y al ajedrez durante cuarenta como escape de la pobreza, de las enfermedades crónicas, de las calumnias y de las injusticias. En su *Liber de ludo aleae* (Libro sobre juegos y azar), publicado póstumamente en 1663, dice que se debe jugar para conseguir el premio de la puesta para compensar el tiempo perdido, y proporciona consejos sobre cómo hacer trampas para asegurar esa compensación.

Después de dedicar su juventud a las matemáticas, la física y el juego, se graduó en medicina por la Universidad de Padua. Practicó este arte, enseñó más tarde en Milán y en Bolonia, y se hizo famoso en toda Europa como médico. También trabajó como profesor de mate-

máticas en varias universidades italianas. En 1570 fue encarcelado por la herejía de realizar el horóscopo de Jesús. Sorprendentemente, el Papa le contrató más tarde como astrólogo. A los setenta y cinco años, poco antes de su muerte en 1576, se jactaba de su fama, de su nieto, de su riqueza, de su conocimiento, de sus amigos poderosos, de su creencia en Dios y de catorce dientes en buen estado.

Sus escritos incluyeron matemáticas, astronomía, astrología, física, medicina y una enorme variedad de otros temas, como aforismos morales (para compensar sus trampas en las cartas). A pesar de su gran preparación en las ciencias, Cardano era un hombre de su tiempo; creía firmemente en la astrología, en los sueños, en los hechizos, en la lectura de la mano, en los portentos y en las supersticiones, y escribió muchos volúmenes sobre estos temas. Es el apologista racional de estas artes ocultas, que, según sostenía, permiten tanta certeza como la navegación o la medicina. También escribió tratados enciclopédicos sobre los habitantes del universo, esto es, sobre los ángeles, demonios y diversas inteligencias, incluyendo en esos libros material robado, sin la menor duda, al distinguido amigo de su padre, Leonardo da Vinci. El material existente ocupa cerca de 7.000 páginas.

La tendencia sincrética, mostrada por los filósofos naturales en sus escritos, que unificaba toda la realidad en trabajos gigantescos, está representada en matemáticas por Cardano. Con diligencia exenta de crítica, extrajo de fuentes antiguas, medievales y contemporáneas todo el conocimiento matemático disponible y lo reunió en una masa enciclopédica, utilizando fuentes tanto teóricas como empíricas. A su acariciada teoría de números, mágica y mística, estaba asociada su predilección por la especulación algebraica, que llevó más lejos que sus contemporáneos. Además de ser un doctor famoso, Cardano se distinguió de los demás filósofos naturales ilustrados del siglo XVI por su más profundo interés en las matemáticas. Pero las matemáticas no eran un método para él; eran un talento mágico especial y una forma de especulación cargada emocionalmente.

Uno de los menos conocidos humanistas, Ignacio Danti (1537-1586), profesor de matemáticas en Bolonia, escribió un libro de matemáticas para profanos que reducía toda la matemática pura y aplicada a series de cuadros sinópticos. *Le scienze matematiche ridotte in tavole* (Bolonia, 1577) es característica del espíritu clasificador de los tiempos; sirvió para dirigir el curso de la instrucción matemática en las escuelas al final del siglo XVI. Danti fue uno de los pocos matemáticos y astrónomos que abogaron por considerar la matemáti-

ca aplicada como una rama del conocimiento (como hizo Galileo más tarde). Los temas cubiertos son significativos porque muestran la variedad de las matemáticas en aquella época: aritmética, geometría, música, astrología, goniometría (especialmente la medida de volúmenes), meteorología, dióptrica, geografía, hidrografía, mecánica, arquitectura, arquitectura militar, pintura y escultura. Los primeros cuatro temas representaban la matemática pura; los restantes, la matemática aplicada.

Esfuerzos humanistas característicos son también evidentes en las investigaciones en mecánica realizadas por prestigiosos matemáticos tales como Guidobalde del Monte (1545-1607), Bernardino Baldi (1553-1617) y Giovanni Battista Benedetti (1530-1590). Estos hombres casi no comprendieron los teoremas de Arquímedes; los trabajos de Pappus tenían más sentido para ellos y tenían un mayor atractivo, porque Pappus había explicado con más detalle las demostraciones dadas en anteriores clásicos griegos. Se desviaron poco de los escolásticos tratando los problemas normales y conocidos, y se limitaron a la corrección de afirmaciones y teoremas aislados. Aceptaron muchas cosas que eran falsas y, además, les faltó la habilidad para separar las ideas vivas y vitales de las agotadas y sin valor. Su formación humanística les inclinó a incorporar en deducciones euclídeas todo el conocimiento antiguo y nuevo, sin tener en cuenta cómo ajustaba todo ello con los experimentos. Como consecuencia, sus facultades de crítica se debilitaron y su propia experiencia perdió valor. Su experimentación no tenía ingredientes mágicos, y su erudición era sobre todo humanística pero, en principio y en esencia, eran los últimos de los medievalistas más que los fundadores de nuevos métodos de pensamiento e investigación. Debido a su dependencia con respecto a los antiguos modos de pensamiento, los matemáticos y físicos italianos Francesco Maurolico (1494-1575), Benedetti, Baldi y del Monte, no aportaron contribuciones innovadoras o revolucionarias en la formulación o resolución de problemas matemáticos o físicos. Estos fueron los hombres a quienes más tarde Galileo llamaría generosamente sus maestros, y quienes, en algunos aspectos, le prepararon el camino.

5. El clamor por la reforma de la ciencia

Como en los siglos anteriores de su existencia, las matemáticas iban a obtener sus principales inspiraciones y temas de las ciencias físicas. Sin embargo, para que la ciencia floreciera era esencial que los europeos rompieran con la adherencia servil a la autoridad. Ya bastantes se habían dado cuenta de que el método de la ciencia debía ser cambiado; iniciaron una verdadera ruptura con el escolasticismo y con la aceptación sin crítica del conocimiento griego.

Uno de los primeros en reclamar un nuevo enfoque del conocimiento fue el famoso artista del Renacimiento Leonardo da Vinci (1452-1519). Increíblemente dotado tanto física como mentalmente, alcanzó grandeza como lingüista, botánico, zoólogo, anatomista, geólogo, músico, escultor, pintor, arquitecto, inventor e ingeniero. Leonardo se distinguió en primer lugar por su desconfianza del conocimiento que muchos estudiosos profesaban tan dogmáticamente. Estos hombres, cuyo conocimiento había sido adquirido sólo en los libros, que él describió como pavoneándose, hinchados y pomposos, adornándose no de sus propios trabajos sino de los de otros, trabajos que ellos se limitaban a repetir, no eran más que los recitadores y trompeteros del conocimiento de otras personas. También criticó los conceptos, métodos y objetivos de los eruditos librescos, porque éstos no se relacionaban con el mundo real. Llegó casi a jactarse de que no era un hombre de letras y de que podía hacer mayores y mejores cosas aprendiendo mediante la experiencia. Y, de hecho, aprendió por sí mismo muchos hechos de las matemáticas, algunos principios de la mecánica y las leyes del equilibrio de la palanca. Realizó observaciones notables sobre el vuelo de los pájaros, el flujo del agua, la estructura de las rocas y la estructura del cuerpo humano. Estudió la luz, el color, las plantas y los animales. Son famosas sus palabras: «Si no te basas en los fundamentos adecuados de la naturaleza, trabajarás con poco honor y menos provecho.» La experiencia, decía, nunca decepciona, aunque nuestro juicio sí pueda hacerlo. «En el estudio de las ciencias que dependen de las matemáticas, los que no consultan a la naturaleza sino a autores no son los hijos de la naturaleza, sino sólo sus nietos.»

Leonardo creía en la combinación de la teoría y la práctica. Dice: «Quien ama la práctica sin la teoría es como el marino que se embarca sin timón ni brújula, y nunca sabe dónde amarrará.» Por otra parte, decía, la teoría sin la práctica no puede sobrevivir, y muere tan

rápidamente como vive. «La teoría es el general; los experimentos son los soldados.» Quería utilizar la teoría para dirigir los experimentos.

Sin embargo, Leonardo no comprendió completamente el verdadero método de la ciencia. De hecho, no tenía ni metodología ni ninguna filosofía subyacente. Su trabajo fue el de un investigador práctico de la naturaleza, motivado por móviles estéticos, pero sin ninguna otra dirección. Estaba interesado en las relaciones cuantitativas y las buscó, y en este aspecto fue un precursor de la ciencia moderna. Sin embargo, no era tan conscientemente cuantitativo como Galileo. Mientras que sus escritos sobre mecánica y ciencia fueron utilizados por hombres del siglo XVI tales como Cardano, Baldi, Tartaglia y Benedetti, no constituyeron ningún estímulo para Galileo, Descartes, Stevin y Roberval.

Los puntos de vista de Leonardo sobre las matemáticas y su conocimiento operativo y utilización del mismo eran peculiares de su tiempo e ilustran su espíritu y enfoque. Leyendo a Leonardo se encuentran muchas afirmaciones que sugieren que era un matemático ilustrado y un filósofo profundo que trabajó al nivel del matemático profesional. Dice, por ejemplo: «El hombre que desacredita la certeza suprema de las matemáticas está alimentando la confusión, y no puede silenciar nunca las contradicciones de las ciencias sofistas, que conducen a la charlatanería inacabable... porque ninguna búsqueda humana puede llamarse ciencia a menos que prosiga su camino a través de la exposición y demostración matemática.» Para ir más allá de la observación y de la experiencia sólo había para él un camino fiable a través de las decepciones y los espejismos —las matemáticas—. Sólo manteniéndose unida a las matemáticas podía la mente penetrar con seguridad en el laberinto del pensamiento intangible e insustancial. La naturaleza actúa mediante leyes matemáticas, y las fuerzas y las operaciones de la naturaleza deben de ser estudiadas mediante investigaciones cuantitativas. Estas leyes matemáticas, a las que uno debe acercarse a través de la experiencia, son el objetivo del estudio de la naturaleza. Sobre la base de tales pronunciamientos, a Leonardo se le considera a menudo, sin ninguna duda, como mejor matemático de lo que realmente era. Cuando se examinan las notas de Leonardo se cae en cuenta de lo poco que sabía de matemáticas, y de que su enfoque era empírico e intuitivo.

Más influyente instando una reforma en los métodos de la ciencia fue Francis Bacon (1561-1626). Bacon buscó métodos para obtener verdades en las esferas intelectual, moral, política y física. Aunque ya

estaban produciéndose cambios en los métodos de las ciencias físicas durante el siglo XVI, el público en general e incluso muchos hombres de letras no eran conscientes de ello. La elocuencia sublime de Bacon, su amplia sabiduría, sus visiones globales y valientes pronunciamientos sobre el futuro hicieron que muchos consideraran con más atención lo que estaba ocurriendo y prestaran atención a lo que él dibujaba como la «Gran Instauración». Cuando la gente se dio cuenta finalmente de que la ciencia estaba comenzando a realizar los avances que Bacon había anunciado, le aclamaron como el autor y dirigente de la revolución que él había, solamente, percibido con anticipación. En realidad, comprendió mejor que sus contemporáneos los cambios que estaban teniendo lugar.

La característica más sobresaliente de su filosofía era el anuncio confiado y enfático de una nueva era en el progreso de la ciencia. En 1605 publicó su tratado *Advancement of Learning;* fue seguido por el *Novum Organum* (Nuevo método) de 1620. En este último libro es más explícito. Señala la debilidad y escasos resultados de los esfuerzos anteriores en el estudio de la naturaleza. La ciencia, dice, se ha desarrollado sólo en medicina y matemáticas, o se ha utilizado para la preparación de jóvenes inmaduros. El progreso está en un cambio del método. Todo conocimiento comienza con la observación. Entonces proporciona su extraordinaria contribución: una insistencia en una «inducción graduada y sucesiva» en lugar de una generalización apresurada. Bacon dice: «Hay dos caminos, y sólo pueden ser dos, para buscar y encontrar la verdad. Uno de ellos, desde los sentidos y los casos particulares, vuela a los axiomas más generales, y desde esos principios y sus verdades, establecidos de una vez por todas, inventa y enjuicia los axiomas intermedios. El otro método acumula axiomas a partir de los sentidos y de los casos particulares, ascendiendo en forma continuada y por grados, de modo que, al final, llega a los axiomas más generales; este último camino es el verdadero, pero no ha sido intentado hasta ahora.» Por «axiomas» él entiende proposiciones generales a las que se llega mediante inducción, y adecuados para servir de punto de partida para un razonamiento deductivo.

Bacon atacó el enfoque escolástico de los fenómenos naturales con estas palabras: «Los axiomas descubiertos mediante discusiones no pueden avalar el descubrimiento de nuevas obras; porque la sutileza de la naturaleza es muchas veces mayor que la sutileza de las discusiones... Los errores radicales en la primera elaboración del pensamiento no pueden subsanarse por la excelencia de su funcionamiento o

correcciones subsiguientes... Debemos conducir a los hombres a los casos particulares, mientras que los mismos hombres, por su parte, deben esforzarse por mantener disponibles sus nociones e ideas y comenzar a familiarizarse con los hechos.»

Bacon no se dio cuenta de que la ciencia debe medir para obtener leyes cuantitativas. En otras palabras, no vio qué tipos de investigaciones graduales se necesitaban y el orden en que debían ser realizadas. Tampoco valoró la inventiva que todo descubrimiento requiere. De hecho, dice que «no se deja mucho a la agudeza y poder del ingenio, pero todos los grados de ingenio e intelecto están al mismo nivel».

Aunque no lo creó, Bacon publicó el manifiesto en favor del método experimental. Atacó los sistemas filosóficos preconcebidos, las creaciones mentales y muestras estacionarias del conocimiento. El trabajo científico, dijo, no debe enredarse en una búsqueda de causas finales, lo cual pertenece a la filosofía. La lógica y la retórica son útiles sólo para organizar lo que ya sabemos. Rodeemos y acerquémonos a la naturaleza, y luchemos mano a mano con ella. No realicemos deshilvanados y fortuitos experimentos; seamos sistemáticos, directos y dirigidos. La matemática tiene que ser una sirvienta de la física. En su conjunto, Bacon ofreció un programa fascinante para las futuras generaciones.

Otra doctrina, y programa, se asocia a Francis Bacon, aunque le antecede. En su conjunto, los griegos se habían contentado con obtener de su ciencia y matemáticas una comprensión de los caminos de la naturaleza. Los pocos científicos del primer período medieval y los escolásticos estudiaron ampliamente la naturaleza para determinar la causa final o el propósito al que se orientaban los fenómenos. Sin embargo, los árabes, un pueblo más práctico, estudiaron la naturaleza para adquirir poder sobre ella. Sus astrólogos, profetas y alquimistas buscaron el elixir de la vida, la piedra filosofal, métodos para convertir metales menos útiles en más útiles, y propiedades mágicas de plantas y animales para prolongar la vida del hombre, curar sus enfermedades y enriquecerle materialmente. Mientras que estas pseudociencias continuaban floreciendo en la época medieval, algunos de los escolásticos más racionales —como, por ejemplo, Robert Grosseteste y Roger Bacon— comenzaron a considerar el mismo propósito, pero mediante investigaciones más propiamente científicas. Gracias a los exhortos de Francis Bacon, la dominación de la naturaleza se convirtió en una doctrina positiva y en una motivación extendida.

Bacon deseaba utilizar el conocimiento. Quería dominar a la naturaleza para que estuviera al servicio del hombre y para su bienestar, y no para placer y deleite de los estudiosos. Como decía, la ciencia tiene que remontarse a los axiomas y desde allí descender a las aplicaciones. En *La Nueva Atlántida* Bacon describe una sociedad de estudiosos dotada de espacio y equipamiento para la adquisición de conocimiento útil. Previó que la ciencia puede proporcionar al hombre «infinitos recursos», «dotar a la vida humana de invenciones y satisfacciones, y facilitar la conveniencia y el agrado del hombre». Estos, decía, son los verdaderos y legítimos fines de la ciencia.

Descartes, en su *Discurso del Método*, se hizo eco de este pensamiento:

> Es posible alcanzar un conocimiento que es muy útil en la vida y, en lugar de esa filosofía especulativa que se enseña en las escuelas, podemos encontrar una filosofía práctica mediante la cual, conociendo la fuerza y la acción del fuego, del agua, del aire, de las estrellas, de los cielos y de todos los demás cuerpos que nos rodean, tan nítidamente como conocemos el oficio de nuestros artesanos, podamos de la misma manera utilizarlos en todos aquellos usos para los que están adaptados y, por tanto, convertirnos en los dominadores y poseedores de la naturaleza.

El químico Robert Boyle dijo: «Las satisfacciones de la humanidad pueden aumentar mucho mediante la aportación a las actividades de la misma de la visión del naturalista.»

El desafío lanzado por Bacon y Descartes fue aceptado rápidamente, y los científicos se sumergieron optimistamente en la tarea de dominar y entender la naturaleza. Estas dos motivaciones son todavía las dos principales fuerzas conductoras y, de hecho, las interconexiones entre la ciencia y la ingeniería crecieron rápidamente desde el siglo XVII en adelante.

Este programa fue adoptado aún más seriamente por los gobiernos. La Academia Francesa de Ciencias, fundada por Colbert en 1666, y la Real Sociedad de Londres, fundada en 1662, se dedicaron al cultivo de «tal conocimiento, en la medida en que tienda a ser útil» y a hacer ciencia «útil a la vez que atractiva».

6. El nacimiento de empirismo

Mientras que los reformadores de la ciencia pedían el retorno a la naturaleza y la necesidad de hechos experimentales, los artesanos de orientación práctica, los ingenieros y los pintores estaban en realidad encontrándose con los duros hechos de la experiencia. Utilizando la visión intuitiva natural de los profanos, y buscando no los significados últimos sino sólo las explicaciones útiles de los fenómenos que encontraban en su trabajo, estos técnicos construyeron un conocimiento que desafiaba, en realidad, las distinciones complicadas, las largas derivaciones etimológicas de los significados, las complicadas argumentaciones lógicas y las pomposas citas de autoridades griegas y romanas de que hacían gala los cultos estudiosos, e incluso los humanistas. Como las realizaciones técnicas de la Europa renacentista fueron superiores y más numerosas que las obtenidas por cualquier otra civilización, el conocimiento empírico adquirido en su obtención fue inmenso.

Los artesanos, ingenieros y artistas tenían que resolver problemas procedentes de ideas mecánicas y propiedades de materiales que tenían que convertir en reales y practicables. Sin embargo, las nuevas perspectivas físicas obtenidas de esta manera fueron impresionantes. Fueron los fabricantes de gafas los que, sin descubrir ninguna ley de la óptica, inventaron el telescopio y el microscopio. Los técnicos llegaron a las leyes acudiendo a los fenómenos. Mediante pasos medidos y graduales, que no había sugerido ninguna visión especulativa del método científico, obtuvieron verdades tan profundas y globales como ninguna conjetura hubiera osado anticipar. Mientras que los reformadores teóricos eran audaces, seguros de sí mismos, impulsivos, ambiciosos, y desdeñosos de la antigüedad, los reformadores prácticos eran cultos, modestos, lentos y receptivos de todo conocimiento, bien derivado de la tradición o de la observación. Ellos trabajaban más que especulaban, trataban con lo particular más que con las generalidades, y aportaban algo a la ciencia, en lugar de definirla o proponer cómo obtenerla. En física, artes plásticas y, en general, los campos de la técnica, la experiencia más que la teoría o la especulación se convirtió en la nueva fuente del conocimiento.

A la vez que el puro empirismo de los artesanos, y en realidad sugerido en parte por los problemas que ellos presentaban, tuvo lugar el aumento gradual de la observación y experimentación sistemáticas, llevadas a cabo por un grupo más culto. Griegos como Aristóteles o

Galeno habían observado mucho y tratado de las inducciones que podrían hacerse sobre la base de las observaciones; pero no se puede decir que los griegos tuvieran nunca una ciencia experimental. La actividad del Renacimiento, muy modesta en extensión, señala el comienzo de la nueva y vasta empresa científica. Los grupos más significativos de los experimentalistas del Renacimiento fueron los fisiólogos, zoólogos y botánicos capitaneados por Andrea Vesalio (1514-1564), Ulises Aldrovandi (1522-1605) y Andre Cesalpino (1519-1603), respectivamente.

En el área de las ciencias físicas, el trabajo experimental de William Gilbert (1540-1603) sobre el magnetismo fue, con mucho, el más extraordinario. En su famosa *De Magnete* (1600), afirma explícitamente que debemos comenzar a partir de los experimentos. Aunque respetaba a los antiguos porque provenía de ellos una corriente de sabiduría, desdeñaba a quienes citaban a otros como autoridades y no experimentaban o verificaban lo que se les había comunicado. Sus series de sencillos experimentos, cuidadosamente realizados y detallados, son un clásico del método experimental. Puntualiza, incidentalmente, que Cardano, en su *De Rerum Varietate*, había descrito una máquina de movimiento perpetuo y comenta: «Condenen los dioses estos trabajos falsos, deshonestos y distorsionados, que no hacen más que confundir las mentes de los estudiantes.»

Hemos estado señalando los abigarrados intereses prácticos que condujeron a un vasto desarrollo del estudio de la naturaleza y el impulso consiguiente a la experimentación sistemática. Junto a este trabajo práctico, con mucha independencia con respecto a él pero sin olvidarlo, algunos hombres continuaron esforzándose en el objetivo, más amplio, de la ciencia-la comprensión de la naturaleza. El trabajo de los últimos escolásticos sobre la caída de los cuerpos, descrito en el capítulo anterior, fue continuado en el siglo XVI. El objetivo predominante era el de obtener las leyes básicas del movimiento. El trabajo sobre el movimiento de proyectiles, a menudo descrito como una respuesta a necesidades prácticas, fue motivado mucho más por amplios intereses científicos en mecánica. El trabajo de Copérnico y Kepler en astronomía (cap. 12, sec. 5) estaba motivado ciertamente por el deseo de mejorar la teoría astronómica. Incluso los artistas del Renacimiento intentaron penetrar en la esencia de la realidad.

Afortunadamente, los técnicos y científicos comenzaron a reconocer intereses comunes y a valorar la ayuda que los unos podían obtener de los otros. Los técnicos de los siglos XV y XVI, los primeros

ingenieros que habían confiado en la destreza manual, ingeniosidad mecánica y pura inventiva y cuidaban menos de los principios, se hicieron conscientes de la ayuda que podían obtener, para la práctica, de la teoría. Por su parte, los científicos se percataron de que los artesanos estaban descubriendo hechos de la naturaleza que una teoría correcta debía explicar, y que podían encontrar en el trabajo de los artesanos prometedoras sugerencias para su investigación. En el párrafo de apertura de su *Diálogos referentes a dos nuevas ciencias* (las ciencias eran la resistencia de materiales y la teoría del movimiento), Galileo reconoce esta inspiración para sus investigaciones. «La constante actividad que vosotros, venecianos, mostráis en vuestro famoso arsenal sugiere a la mente estudiosa un amplio campo para la investigación, especialmente esa parte del trabajo que se relaciona con la mecánica; porque en este departamento todos los tipos de instrumentos y máquinas están siendo construidos constantemente por muchos artesanos, entre quienes debe de haber algunos que, en parte por experiencias heredadas y en parte por su propia observación, han llegado a ser muy hábiles y expertos en las explicaciones.»

Los intereses prácticos y puramente científicos se fundieron en el siglo XVII. Cuando los principios generales y los problemas surgieron de las necesidades empíricas, y el conocimiento matemático de los griegos había llegado a ser completamente accesible para los científicos, éstos estuvieron más capacitados para continuar desarrollando la ciencia pura. Sin perder de vista el objetivo de entender el plan del universo, también intentaron voluntariamente impulsar la práctica. El resultado fue un desarrollo de la actividad científica en una escala sin precedentes, además de mejoras técnicas de peso y de gran alcance que culminaron en la Revolución Industrial.

La gran importancia de los comienzos de la ciencia moderna para nosotros es, por supuesto, que preparó el camino para los principales desarrollos en matemáticas. Su efecto inmediato fue su relación con problemas concretos. Puesto que los matemáticos del Renacimiento trabajaron para las repúblicas y los príncipes, y colaboraron con los arquitectos y obreros manuales —Maurolico fue un ingeniero que trabajaba para la ciudad de Messina, Baldi fue un matemático que trabajaba para el Duque de Urbino, Benedetti fue el ingeniero jefe del Duque de Saboya, y Galileo fue matemático de corte del Gran Duque de Toscana— tuvieron en cuenta las observaciones y las experiencias de la gente práctica. Hasta la época de Galileo, el impacto de los técnicos y de los arquitectos puede verse ampliamente en el trabajo de

Nicolo Tartaglia (1499?-1557), un genio que fue autodidacta en la ciencia de su tiempo. Tartaglia realizó la transición del matemático práctico al ilustrado, separando con criterio problemas y observaciones útiles del conocimiento empírico. Su singularidad radica en su realización y en su completa independencia de las influencias mágicas que caracterizan el trabajo de su rival Cardano. La posición de Tartaglia está a mitad de camino entre Leonardo y Galileo —no sólo cronológicamente, sino porque su trabajo sobre las matemáticas de problemas dinámicos elevaron ese tema a la categoría de nueva ciencia e influyó en los predecesores de Galileo.

El efecto a largo plazo fue que la matemática moderna, guiada por la doctrina platónica de que es la esencia de la realidad, creció casi enteramente a partir de los problemas de la ciencia. Bajo la nueva línea directriz de estudiar la naturaleza y obtener leyes que englobaran observaciones y hechos experimentales, las matemáticas rompieron con la filosofía y se unieron a las ciencias físicas. Las consecuencias para las matemáticas fueron una explosión de actividad y de creación original que fue la más prolífica en su historia.

Bibliografía

Ball, W. W. R.: *A Short Account of the History of Mathematics,* Dover (reimpresión), 1960, caps. 12-13.

Burtt, E. A.: *Los fundamentos metafísicos de la Ciencia Moderna,* Buenos Aires, Suramericana, 1945.

Butterfield, Herbert: *Los Orígenes de la Ciencia Moderna,* Madrid, Taurus, 1982.

Cajori, Florian: *A History of Mathematics,* 2nd ed., Macmillan, 1919, pp. 128-145.

Cardano, Gerolamo: *Opera Omnia,* Johnson Reprint Corp., 1964.

—: *Mi vida,* Madrid, Alianza, 1991.

—: *The Book on Game of Chance,* Holt, Rinehart and Winston, 1961.

Claggett, Marshall, ed.: *Critical Problems in the History of Science,* University of Wisconsin Press, 1959, pp. 3-196.

Crombie, A. C.: *Historia de la ciencia de Agustín a Galileo,* Madrid, Alianza, 1979.

—: *Robert Grosseteste and the Origins of Experimental Science,* Oxford University Press, 1953, cap. 11.

Dampier-Whetham, W. C. D.: *Historia de la Ciencia,* Madrid, Tecnos, 1986.

Farrington, B.: *Francis Bacon, filósofo de la revolución industrial,* Madrid, Ayuso, 1971.

Mason, S. F.: *Historia de las Ciencias*, Madrid, Alianza, 1988.
Ore, O.: *Cardano: The Gambling Scholar*, Princeton University Press, 1953.
Randall, John H., Jr.: *The Making of the Modern Mind*, Houghton Mifflin, 1940, caps. 6, 9.
Russell, Bertrand: *Historia de la filosofía occidental*, Madrid, Espasa-Calpe, 1984.
Smith, David Eugene: *History of Mathematics*, Dover (reimpresión), 1958, vol. I, pp. 242-265, y cap. 8.
Smith, Preserved: *A History of Modern Culture*, Holt, Rinehart and Winston, 1940, vol. I, caps. 5-6.
Strong, Edward, W.: *Procedures and Metaphysics*, University of California Press, 1936; reimpreso por Georg Olms, 1966, pp. 1-134.
Taton, René, ed.: *The Beginnings of Modern Science*, Basic Books, 1964, pp. 3-51, y 82-177.
Vallentin, Antonina: *Leonardo da Vinci*, Viking Press, 1938.
White, Andrew D.: *Historia del conflicto entre Ciencia y Teología*, Siglo XXI, 1972.

Capítulo 12
LAS CONTRIBUCIONES MATEMATICAS EN EL RENACIMIENTO

> El principal propósito de todas las investigaciones sobre el mundo exterior debe ser descubrir el orden y la armonía racionales que han sido impuestos por Dios y que Él nos ha revelado en el lenguaje de las matemáticas.
>
> JOHANNES KEPLER

1. Perspectiva

Aunque los hombres del Renacimiento sólo percibieron ligeramente las perspectivas, valores y objetivos de las obras griegas, dieron algunos pasos originales en matemáticas y realizaron avances en otros campos, que prepararon el camino para el extraordinario desarrollo de nuestro tema en el siglo XVII.

Los artistas fueron los primeros en manifestar la renovación del interés en la naturaleza y en aplicar seriamente la doctrina griega de que las matemáticas son la esencia de la realidad de la naturaleza. Los artistas eran autodidactas y aprendían mediante la práctica. Les llegaron filtrados algunos fragmentos del conocimiento griego pero, en su conjunto, sintieron más que percibieron las ideas y perspectivas intelectuales griegas. Hasta cierto punto, esto fue una ventaja porque, careciendo de una escolarización formal, estaban libres de todo adoctrinamiento. Además, disfrutaron de libertad de expresión, porque su trabajo se consideraba «innocuo».

Los artistas del Renacimiento eran, por su profesión, hombres universales, esto es, eran contratados por los príncipes para realizar todo tipo de tareas, desde la creación de grandes pinturas hasta el

diseño de fortificaciones, canales, puentes, máquinas de guerra, palacios, edificios públicos e iglesias. En consecuencia, estaban obligados a aprender matemáticas, física, arquitectura, ingeniería, tallado de las piedras, el trabajo de los metales, anatomía, el trabajo de la madera, óptica, estática e hidráulica. Realizaron trabajos manuales, pero también se ocuparon de los problemas más abstractos. En el siglo XV, al menos, ellos eran los mejores físicos matemáticos.

Para valorar sus contribuciones a la geometría, debemos hacer notar sus nuevos objetivos en la pintura. En el período medieval, la glorificación de Dios y la ilustración de los temas bíblicos eran los fines de la pintura. Los fondos dorados sugerían que las personas y objetos retratados estaban en alguna región celestial. Además, se pretendía que las figuras fueran simbólicas más que realistas. Los pintores producían formas planas y sin naturalidad, y no se desviaban de los cauces establecidos. En el Renacimiento, la descripción del mundo real se convirtió en el objetivo de la pintura. Por lo tanto, los artistas emprendieron el estudio de la naturaleza para reproducirla fielmente en sus lienzos, y se enfrentaron al problema matemático de presentar el mundo real tridimensional en un lienzo bidimensional.

Filippo Brunelleschi (1377-1446) fue el primer artista que estudió y utilizó intensivamente las matemáticas. Giorgio Vasari (1511-1574), el artista y biógrafo italiano, dice que el interés de Brunelleschi en las matemáticas le llevó a él a estudiar perspectiva, y que empezó a pintar para aplicar la geometría. Leyó a Euclides, a Hiparco y a Vitello en matemáticas y óptica, y aprendió matemáticas del matemático florentino Paolo del Pozzo Toscanelli (1397-1482). Los pintores Paolo Ucello (1397-1475) y Masaccio (1401-1428) también se interesaron por principios matemáticos para desarrollar un sistema de perspectiva realista.

El genio teórico en la perspectiva matemática fue Leone Battista Alberti (1404-1472), quien presentó sus ideas en *Della Pittura* (1435), impreso en 1511. Este libro, profundamente matemático en su carácter, también incluye algo de óptica. Su otro importante trabajo matemático es *Ludi Mathematici* (1450), que contiene aplicaciones a la mecánica, agrimensura, cálculo del tiempo y fuego de artillería. Alberti concibió el principio que se convertiría en la base del sistema matemático de perspectiva adoptado y perfeccionado por sus sucesores artistas. Propuso pintar lo que ve un ojo, aunque era muy consciente de que en la visión normal ambos ojos ven la misma escena desde posiciones ligeramente distintas, y que sólo a través de la

reconciliación de las dos imágenes en el cerebro se percibe la profundidad. Su plan era obtener la ilusión de la profundidad mediante instrumentos tales como la luz y la sombra, y la disminución del color con la distancia. Su principio básico puede explicarse en los siguientes términos. Entre el ojo y la escena interponía una pantalla de vidrio en posición vertical. Entonces imaginaba líneas de luz desde el ojo o punto de estación hasta cada punto de la escena misma. Llamaba a estas líneas una pirámide de rayos o una proyección. Donde estos rayos atravesaban la pantalla de vidrio (la imagen plana), imaginaba puntos marcados; a esta colección de puntos la llamaba una sección. El hecho significativo sobre esto es que creaba la misma impresión sobre el ojo que la escena misma, porque de la sección provienen las mismas líneas de luz que de la escena original. En consecuencia, el problema de pintar en forma realista es el de obtener una sección verdadera sobre la pantalla de vidrio o, en la práctica, sobre un lienzo. Por supuesto que la sección depende de la posición del ojo y de la de la pantalla. Esto significa sólo que pueden hacerse diferentes pinturas de la misma escena.

Como el pintor no mira a través de su lienzo para determinar la sección, debe disponer de reglas basadas en teoremas matemáticos que le digan cómo dibujarla. Alberti proporcionó algunas reglas [1] correctas en su *Della Pittura*, pero no dio todos los detalles. Pretendió que su libro fuera como un sumario que se completara con discusiones con sus compañeros pintores y, de hecho, se disculpa por su brevedad. Intentó hacer su material concreto, más que formal y riguroso, y por ello dio teoremas y construcciones sin demostraciones.

Además de presentar los conceptos de proyección y sección, Alberti suscitó una cuestión muy significativa. Si se interponen dos pantallas de vidrio entre el ojo y la escena, las secciones en ellas serán diferentes. Además, si el ojo mira la misma escena desde dos posiciones distintas y, en cada caso, se interpone una pantalla de vidrio entre el ojo y la escena, otra vez la sección será diferente. Y, sin embargo, la sección corresponde a la misma figura. Por lo tanto, deben tener algunas propiedades en común. La cuestión es: ¿cuál es la relación matemática entre dos cualesquiera de estas secciones o cuáles son las propiedades matemáticas que tienen en común? Esta cuestión se

---

[1] Para algunas de estas reglas y pinturas construidas de acuerdo con ellas, ver la obra del autor, *Mathematics in the Western Culture*, Oxford University Press, 1953.

convirtió en el punto de partida del desarrollo de la geometría proyectiva.

Aunque bastantes artistas escribieron libros sobre perspectiva matemática y compartían la filosofía del arte de Alberti, podemos mencionar aquí sólo uno o dos de los principales. Leonardo creía que la pintura debe ser una reproducción exacta de la realidad, y que la perspectiva matemática lo permitiría. Era «el timón e hilo conductor de la pintura» e incluía la óptica aplicada y la geometría. La pintura, para él, era una ciencia porque revela la realidad en la naturaleza; por esta razón es superior a la poesía, a la música y a la arquitectura. Los escritos de Leonardo sobre la perspectiva están contenidos en su *Tratatto della pittura* (1651), compilado por algún autor desconocido que utilizó lo más valioso de las notas de Leonardo sobre el asunto.

El pintor que estableció los principios matemáticos de la perspectiva en una forma bastante completa fue Piero della Francesca (*c.* 1410-1492). También él consideraba la perspectiva como la ciencia de la pintura y quiso corregir y extender el conocimiento empírico a través de las matemáticas. Su trabajo principal, *De prospettiva pingendi* (1482-1487), aportó algunos avances a las ideas de Alberti de proyección y sección. En general, suministra procedimientos útiles para los artistas, y sus indicaciones incluyen el empleo de bandas de papel, madera y cosas análogas. Para ayudar al artista, le proporciona, como Alberti, definiciones inteligibles intuitivamente. Entonces le brinda teoremas que «demuestra» mediante construcciones o por un cálculo aritmético de razones. Fue el pintor matemático y el artista científico por excelencia, y sus contemporáneos así le consideraron. Fue también el mejor geómetra de su tiempo.

Sin embargo, de todos los artistas del Renacimiento, el mejor matemático fue el alemán Albrecht Dürer (Alberto Durero) (1471-1528). Su *Underweysung der Messung mid dem Zyrkel und Rychtscheyd* (Instrucción en la medida con regla y compás, 1525), un libro de geometría, sobre todo, fue realizado para transmitir a los alemanes el conocimiento que Durero había adquirido en Italia y, en particular, para ayudar a los artistas con la perspectiva. El libro, más preocupado por la práctica que por la teoría, fue muy influyente.

La teoría de la perspectiva se enseñaba en las escuelas de pintura desde el siglo XVI en adelante, de acuerdo con los principios establecidos por los maestros de los que nos hemos ocupado. Sin embargo, sus tratados sobre perspectiva habían consistido, en su conjunto, en preceptos, reglas y procedimientos *ad hoc:* les había faltado una sólida

base matemática. En el período comprendido entre 1500 y 1600, los artistas y matemáticos siguientes situaron el tema sobre una base deductiva satisfactoria, y pasó de ser un arte casi empírico a una verdadera ciencia. Trabajos definitivos sobre perspectiva fueron escritos mucho más tarde por los matemáticos del siglo XVIII Brook Taylor y J. H. Lambert.

2. La geometría propiamente dicha

Aparte de la perspectiva, los desarrollos que experimentó la geometría durante los siglos XV y XVI no fueron muy impresionantes. Uno de los temas de que trataron Durero, Leonardo y Luca Pacioli (*c.* 1445-*c.* 1514), un monje italiano alumno de Piero della Francesca y amigo y profesor de Leonardo, fue el de la inscripción de polígonos regulares en circunferencias. Intentaron tales construcciones mediante una regla y un compás de apertura fija, una limitación ya considerada por el árabe Abû'l-Wefâ, pero sólo proporcionaron métodos aproximados.

La construcción del pentágono regular era un problema de gran interés, porque surgía en el diseño de fortificaciones. En los *Elementos* (libro IV, proposición 11), Euclides había dado una construcción no limitada por un compás de apertura fija. El problema de dar una construcción exacta con esta limitación fue tratado por Tartaglia, Ferrari, Cardano, del Monte, Benedetti y muchos otros matemáticos del siglo XVI. Benedetti, entonces, amplió el problema y se propuso resolver todas las construcciones euclídeas con una regla y un compás de apertura fija. El problema general fue resuelto por el danés George Mohr (1640-1697) en su *Compendium Euclidis Curiosi* (1673).

Mohr demostró también, en su *Euclides Danicus* (1672), que las construcciones que podían realizarse con una regla y un compás podían realizarse también con sólo un compás. Por supuesto que sin la regla no es posible dibujar la línea recta que une dos puntos, pero dados dos puntos se pueden construir los puntos de intersección de la recta y la circunferencia, y dados dos pares de puntos, se puede construir el punto de intersección de las dos rectas determinadas por los dos pares. El hecho de que sólo un compás baste para realizar las construcciones euclídeas fue redescubierto por Lorenzo Mascheroni (1750-1800) y publicado en su *La geometria del compasso.*

Otro tema de interés para los griegos, el de los centros de gravedad

de los cuerpos, también fue considerado por los geómetras del Renacimiento. Leonardo, por ejemplo, dio un método correcto y uno incorrecto para obtener el centro de gravedad de un trapezoide isósceles. Dio después, sin demostración, la situación del centro de gravedad del tetraedro, a la cuarta parte del segmento que une el centro de gravedad de la base triangular con el vértice opuesto, contada desde la base.

Dos ideas geométricas originales aparecen en trabajos menores de Durero. La primera de ellas es la de considerar curvas en el espacio. Comienza estudiando curvas del espacio helicoidales y considera la proyección de estas curvas sobre el plano. Las proyecciones son varios tipos de espirales, y Durero muestra cómo construirlas. También presenta la epicicloide, que es el lugar geométrico descrito por un punto determinado de una circunferencia que gira apoyándose en el exterior de otra circunferencia fija. La segunda idea es la proyección ortogonal de curvas y de figuras humanas en dos y tres planos mutuamente perpendiculares. Esta idea, que Durero sólo rozó, fue desarrollada a finales del siglo XVIII, dentro de la geometría descriptiva, por Gaspard Monge.

El trabajo de Leonardo, Piero, Pacioli y Durero en geometría pura no es, ciertamente, significativo desde el punto de vista de la obtención de resultados nuevos. Su valor principal fue el de que difundió ampliamente algunos conocimientos de geometría, aunque fueran algo bastos comparados con los que alcanzaron los griegos. La cuarta parte de la obra de Durero *Underweysung*, junto con la de Piero *De Corporibus Regularibus* (1487) y la de Pacioli *De Divina Proportione* (1509) renovaron el interés por la estereometría (medida de figuras sólidas), que floreció en la época de Kepler.

Otra actividad geométrica, la elaboración de mapas, sirvió para estimular posteriores investigaciones geométricas. Las exploraciones geográficas habían revelado la inadecuación de los mapas existentes; al mismo tiempo, se iban desvelando nuevos conocimientos geográficos. La confección e impresión de mapas había comenzado en la segunda mitad del siglo XV en centros como Amberes y Amsterdam.

El problema de la confección de mapas surge del hecho de que una esfera no puede cortarse, abrirse y extenderse sobre un plano sin distorsionar las distancias. Además, las direcciones (ángulos) o áreas, o ambos, pueden distorsionarse también. El método nuevo más significativo de confección de mapas se debe a Gerard Kremer, conocido también como Mercator (1512-1594), quien dedicó su vida a la

ciencia. En 1569 publicó un mapa utilizando la famosa proyección de Mercator. En este esquema, las líneas de latitud y de longitud son rectas. Las líneas de longitud están igualmente espaciadas, pero el espaciado entre las líneas de latitud se incrementa. El propósito de este incremento es el de mantener correcto el cociente entre el largo de un minuto de longitud y el de un minuto de latitud. Sobre la esfera, un cambio de 1′ de latitud corresponde a 6087 pies; pero un cambio de 1′ de longitud sólo es igual a 6087 pies en el ecuador. Por ejemplo, en la latitud 20°, un cambio de 1′ en longitud equivale a 5722 pies, proporcionando el cociente

$$\frac{1' \text{ de cambio en longitud}}{1' \text{ de cambio en latitud}} = \frac{5722}{6087}.$$

Para que este cociente se mantenga en el mapa de líneas rectas de Mercator, en el que las líneas de longitud están igualmente espaciadas y cada minuto de cambio equivale a 6087 pies, él aumenta los espacios entre las líneas de latitud mediante el factor $1/\cos L$, a medida que la latitud L aumenta. A 20° de latitud en su mapa un cambio de latitud de 1′ equivale a una distancia de $6087(1/\cos L)$, es decir, 6450 pies. Por tanto, a 20° de latitud

$$\frac{1' \text{ de cambio en longitud}}{1' \text{ de cambio en latitud}} = \frac{6087}{6450}.$$

y este cociente es igual al cociente 5722/6087.

El mapa de Mercator tiene varias ventajas. Sólo en esta proyección dos puntos del mapa están en el rumbo correcto de la brújula uno con respecto a otro. Por tanto, una curva sobre la esfera cuyo rumbo es la brújula constante, esto es, una curva llamada loxodroma o línea de rumbo, que corta a todos los meridianos según un mismo ángulo, se convierte en una línea recta en el mapa. Las distancias y las áreas no se conservan; de hecho, el mapa distorsiona mucho en los polos. Sin embargo, como la dirección se conserva, también se conserva el ángulo de dos direcciones en un punto, y se dice que el mapa es conforme.

Aunque ninguna idea matemática grande surgió de los trabajos de elaboración de mapas en el siglo XVI, el problema se volvió a

considerar más tarde por otros matemáticos, y eso les llevó a trabajar en geometría diferencial.

3. Algebra

Hasta la aparición de la *Ars Magna* (1545) de Cardano, de la que trataremos en el próximo capítulo, no hubo en el Renacimiento desarrollos trascendentes en álgebra. Sin embargo, vale la pena mencionar el trabajo de Pacioli. Como muchos otros de su siglo, es el conocimiento sistemático más amplio y que se puede aplicar a la vida práctica y espiritual de todo el mundo. También se dio cuenta de las ventajas del conocimiento teórico para realizar el trabajo práctico. La teoría debe dominar y guiar, dice a los matemáticos y técnicos. Como Cardano, pertenecía al círculo humanista. La principal publicación de Pacioli es la *Summa de Arithmetica, Geometria, Proportione et Proportionalita* (1494). La *Summa* era un compendio del conocimiento disponible, y era representativa de su época porque vinculaba las matemáticas con una gran cantidad de aplicaciones prácticas.

El contenido incluía los símbolos numéricos indo-arábigos, que ya se utilizaban en Europa, la aritmética de los negocios, en particular la contabilidad, el álgebra creada hasta entonces, un pobre resumen de los *Elementos* de Euclides y algo de trigonometría tomada de Ptolomeo. La aplicación del concepto de proporción para descubrir el plan en todas las fases de la naturaleza y en el universo mismo era un tema importante. Pacioli llamaba a la proporción «madre» y «reina» y la aplicaba a los tamaños de las partes del cuerpo humano, a la perspectiva e incluso a las mezclas de colores. Su álgebra es retórica; sigue a Leonardo y a los árabes al llamar a la incógnita la «cosa». Al cuadrado de la incógnita Pacioli le llama *census,* que a veces abrevia como *ce* o *Z*; el cubo de la incógnita, *cuba*, se representa a veces como *cu* o *C*. También aparecen otras abreviaturas para palabras, tales como *p* para más y *ae* para *aequalis*. Escribiendo ecuaciones, cuyos coeficientes son siempre numéricos, coloca los términos en el lado que permita la utilización de coeficientes positivos. Aunque aparezca la sustracción ocasional de un término, por ejemplo, $-3x$, no se utilizan números puramente negativos; sólo se dan las raíces positivas. Utilizó el álgebra para calcular cantidades geométricas, de la misma manera que nosotros utilizaríamos una proporción aritmética para relacionar las longitudes de los lados de dos triángulos semejantes y, quizá, para

obtener una longitud desconocida, aunque la utilización de Pacioli es, a menudo, más complicada. Termina el libro con la observación de que la solución de $x^3 + mx = n$ y de $x^3 + n = mx$ (utilizamos la notación moderna) son tan imposibles como la cuadratura del círculo.

Aunque no había nada de original en la *Summa*, este libro y su *De Divina Proportione* fueron valiosos porque contenían mucho más de lo que se enseñaba en las universidades. Pacioli hizo de intermediario entre lo que existía en los trabajos académicos y el conocimiento adquirido por artistas y técnicos, a quienes intentó ayudar a aprender y a utilizar las matemáticas. Sin embargo, un comentario significativo sobre el desarrollo matemático de la aritmética y el álgebra entre 1200 y 1500 es que la *Summa* de Pacioli, que apareció en 1494, no contenía prácticamente nada más que el *Liber Abaci* de Leonardo de Pisa, de 1202. De hecho, la aritmética y el álgebra de la *Summa* estaban basadas en el libro de Leonardo.

4. Trigonometría

Hasta 1450, la trigonometría era sobre todo trigonometría esférica; la agrimensura continuaba utilizando los métodos geométricos de los romanos. Aproximadamente hacia esa fecha, la trigonometría plana comenzó a tener importancia en agrimensura, aunque Leonardo de Pisa ya había iniciado el método en su *Practica Geometriae* (1220).

Los alemanes llevaron a cabo nuevos trabajos en trigonometría a finales del siglo XV y principios del XVI. Habitualmente, estudiaban en Italia y luego volvían a sus ciudades de origen. En esa época, Alemania se había hecho próspera. Alguna de esa riqueza había sido adquirida a través de la Liga Hanseática del norte de Alemania, que controlaba buena parte del comercio; por ello algunos comerciantes importantes pudieron apoyar económicamente el trabajo de muchos de los que mencionaremos más adelante. El trabajo sobre trigonometría fue motivado por la navegación, el cálculo del calendario y la astronomía, por la que había crecido el interés con la creación de la teoría heliocéntrica, sobre la que trataremos más tarde.

George Peurbach (1423-1461), de Viena, comenzó a corregir las traducciones latinas del *Almagesto*, que se habían hecho de las versiones árabes, pero que él propuso hacer del original griego. También comenzó a hacer tablas trigonométricas más precisas. Sin embargo, Peurbach murió joven, y su trabajo fue continuado por su discípulo

Johannes Müller (1436-1476), conocido como Regiomontano, quien revitalizó la trigonometría en Europa. Habiendo estudiado astronomía y trigonometría con Peurbach en Viena, Regiomontano fue a Roma, estudió griego con el cardenal Bessarion (*c.* 1400-1472), y reunió los manuscritos griegos de los eruditos griegos que habían huido de los turcos. En 1471 se estableció en Nuremberg bajo el patronazgo de Bernard Walther. Regiomontano hizo traducciones de trabajos griegos —las *Secciones Cónicas* de Apolonio y partes de Arquímedes y Herón— y fundó su propia imprenta para imprimirlas.

Siguiendo a Peurbach, adoptó el seno hindú, esto es, la semicuerda del semiarco, y construyó una tabla de senos basada en un radio de 600.000 unidades y otra basada en un radio de 10.000.000 de unidades. También calculó una tabla de tangentes. En la *Tabulae Directionum* (escrita entre 1464 y 1467), dio tablas de tangentes de cinco cifras y subdivisiones decimales de los ángulos, un procedimiento muy poco habitual para aquellos tiempos.

Entre los muchos que construyeron tablas en los siglos XV y XVI, puede mencionarse a George Joachim Rhaeticus (1514-1576), Copérnico, François Vieta (1540-1603) y Bartolomäus Pitiscus (1561-1613). Una característica de su trabajo fue la utilización de radios de cada vez mayor número de unidades, de manera que los valores de las cantidades trigonométricas podían obtenerse en forma más precisa, sin necesidad de utilizar fracciones o decimales. Por ejemplo, Rhaeticus calculó una tabla de senos basada en un radio de 10^{10} unidades y otra basada en uno de 10^{15} unidades, y dio valores para cada 10 segundos de arco. Pitiscus en su *Thesaurus* (1613) corrigió y publicó la segunda tabla de Rhaeticus. La palabra «trigonometría» es suya.

Más fundamental fue el trabajo sobre la resolución de triángulos planos y esféricos. Hasta aproximadamente 1450, la trigonometría esférica consistía en unas reglas sueltas basadas en versiones griegas, hindúes y árabes, la última de las cuales vino de España. Los trabajos de los árabes orientales Abû'l-Wefâ y Nâssir-Eddin no se conocieron en Europa hasta entonces. Regiomontano pudo aprovechar el trabajo de Nâssir-Eddin y, en *De Triangulis,* escrito entre 1462 y 1463, reunió en una forma más efectiva el conocimiento disponible en trigonometría plana, geometría esférica y trigonometría esférica. Obtuvo la ley de los senos para triángulos esféricos, es decir

$$\frac{\operatorname{sen} a}{\operatorname{sen} A} = \frac{\operatorname{sen} b}{\operatorname{sen} B} = \frac{\operatorname{sen} c}{\operatorname{sen} C}$$

y la ley de los cosenos que relaciona los lados, esto es,

$$\cos a = \cos b \ \cos c + \operatorname{sen} b \operatorname{sen} c \cos A.$$

El *De Triangulis* no fue publicado hasta 1533; mientras tanto, Johann Werner (1468-1528) mejoró y publicó las ideas de Regiomontano en *De Triangulis Sphaericis* (1514).

Durante muchos años después del trabajo de Regiomontano, la trigonometría esférica continuó siendo confusa por la necesidad de una multitud de fórmulas, en parte porque Regiomontano en su *De Triangulis*, e incluso Copérnico un siglo después, utilizaron sólo las funciones seno y coseno. Además, los valores negativos para las funciones coseno y tangente de los ángulos obtusos no eran considerados números.

Rhaeticus, que era un discípulo de Copérnico, cambio el significado del seno. En vez de llamar a *AB* (fig. 12.1) el seno de *AD*, llamó a *AB* el seno del ángulo *AOB*. Sin embargo, la longitud de *AB* seguía expresándose en una cantidad de unidades que dependía de la cantidad de unidades elegida como longitud del radio. Como consecuencia del cambio de Rhaeticus, el triángulo *OAB* se convirtió en la estructura básica, y la circunferencia de radio *OA* en algo accesorio. Rhaeticus utilizó las seis funciones.

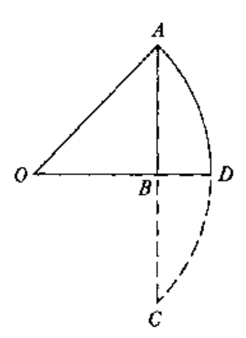


FIGURA 12.1

Las trigonometrías plana y esférica fueron, más tarde, sistematizadas y extendidas ligeramente por François Vieta, abogado de profesión, pero valorado mucho más como el matemático más importante

del siglo XVI. Su *Canon Mathematicus* (1579) fue el primero de sus muchos trabajos sobre trigonometría. En él reunió las fórmulas para la resolución de triángulos planos rectos y oblicuos, e incluyó su propia contribución, la ley de las tangentes:

$$\frac{a-b}{a+b}=\frac{\tan\left(\dfrac{A-B}{2}\right)}{\tan\left(\dfrac{A+B}{2}\right)}.$$

Para triángulos esféricos rectos proporcionó el conjunto completo de fórmulas que se necesitan para calcular cualquier elemento en términos de otros dos cualesquiera, y la regla para recordar esta colección de fórmulas, que ahora llamamos la regla de Napier. También aportó la regla de los cosenos que relaciona los ángulos de un triángulo esférico oblicuo:

$$\cos A = -\cos B \cos C + \operatorname{sen} B \operatorname{sen} C \cos a.$$

Muchas identidades trigonométricas habían sido establecidas por Ptolomeo; Vieta añadió algunas otras. Por ejemplo, obtuvo la identidad

$$\operatorname{sen} A - \operatorname{sen} B = 2 \cos \frac{A+B}{2} \operatorname{sen} \frac{A-B}{2}$$

e identidades para sen $n\theta$ y cos $n\theta$ en términos de sen θ y de cos θ. Estas últimas identidades están contenidas en su *Sectiones Angulares*, publicación póstuma de 1615 [2]. Vieta expresó las identidades y trabajó con ellas en forma algebraica, aunque la notación no era en absoluto moderna.

Utilizó la fórmula del sen $n\theta$ para resolver el problema propuesto por el matemático belga Adrianus Romanus (1561-1615) en su libro *Ideae Mathematicae* (1593) como desafío a todos los franceses. El problema consistía en resolver una ecuación de grado cuarenta y cinco en x. Enrique IV de Francia llamó a Vieta, quien se dio cuenta de que

[2] *Opera*, Leyden, 1646, 287-304.

el problema equivalía a lo siguiente: dada la cuerda de un arco, obtener la cuerda de la cuadragésima quinta parte de ese arco. Ello es equivalente a expresar sen $45A$ en términos de sen A y obtener sen A. Si x = sen A, entonces la ecuación algebraica es de grado cuarenta y cinco en x. Vieta sabía que este problema se podía resolver separando la ecuación en una ecuación de quinto grado y en dos de tercer grado, que podía resolver rápidamente. Dio las 23 raíces positivas, pero ignoró las negativas. En su *Responsum* (1595) [3] explicó su método de resolución.

En el siglo XVI, la trigonometría comenzó a separarse de la astronomía y adquirió el rango de rama de las matemáticas. Siguió siendo importante la aplicación a la astronomía, pero se desarrollaron otras como, por ejemplo, la agrimensura, que garantizaron el estudio del tema desde un punto de vista más independiente.

5. Los principales progresos científicos del Renacimiento

Los matemáticos del Renacimiento prepararon el terreno para el resurgir del estudio matemático en Europa mediante las traducciones de los trabajos griegos y árabes y los trabajos enciclopédicos de compilación del conocimiento existente. Pero las motivaciones y direcciones de las creaciones matemáticas siguientes de los europeos surgieron principalmente de los problemas científicos y tecnológicos, aunque hubo algunas excepciones. El crecimiento del álgebra fue, al menos al comienzo, una continuación de las líneas árabes de actividad, y algunos de los nuevos trabajos en geometría fueron sugeridos por problemas propuestos por artistas.

Con mucho, el desarrollo renacentista más significativo en la motivación de las matemáticas de los dos siglos siguientes fue la revolución en astronomía, que capitanearon Copérnico y Kepler. Cuando los textos griegos estuvieron disponibles, después de, aproximadamente, 1200, tanto la teoría astronómica de Aristóteles (una modificación de la de Eudoxo) como la teoría de Ptolomeo se difundieron ampliamente y se tendió a considerarlas contrapuestas. Estrictamente hablando, tanto los árabes como los últimos astrónomos medievales habían hecho aportaciones para mejorar la precisión de ambos esquemas o para adaptar el esquema de Aristóteles a la

---

[3] *Opera*, 305-324.

teología cristiana. El esquema de Ptolomeo, razonamiento preciso para su época, era puramente matemático y considerado, por tanto, sólo como una hipótesis, y no como una descripción de estructuras reales. La teoría de Aristóteles era la más generalmente aceptada, aunque la de Ptolomeo era más útil para las predicciones astronómicas, la navegación y el cálculo del calendario.

Algunas figuras árabes, del final del medievo y del Renacimiento, como al-Bîrûni (973-1048), Oresme y el cardenal Nicolás de Cusa (1401-1464), quizá en respuesta a ideas griegas, consideraron seriamente que la Tierra pudiera estar girando, y que podría ser igualmente posible construir una teoría astronómica basada en el movimiento de la Tierra alrededor del Sol, pero ninguno desarrolló una nueva teoría.

Entre los astrónomos, Nicolás Copérnico apareció de pronto como un coloso. Nacido en Thorn, Polonia, en 1473, Copérnico estudió matemáticas y ciencia en la universidad de Cracovia. A la edad de veintitrés años fue a Bolonia para proseguir sus estudios, y allí se familiarizó con la doctrina pitagórica y otras doctrinas griegas, en particular con la teoría astronómica. También estudió medicina y derecho canónico. En 1512 volvió a Polonia y se hizo canónigo en la Catedral de Frauenberg, donde permaneció hasta su muerte en 1543. Mientras realizaba sus tareas se dedicó a estudios y observaciones intensivos, que culminaron en una teoría astronómica revolucionaria. Esta realización en el terreno del pensamiento sobrepasa decisivamente en significación, valentía y magnificencia a la conquista de los mares.

Es difícil determinar cuál fue la causa de que Copérnico derrocara la teoría de Ptolomeo, de mil cuatrocientos años de antigüedad. Las indicaciones en el prefacio de su obra clásica, *De Revolutionibus Orbium Coelestium* (Sobre las revoluciones de las esferas celestes, 1543) son incompletas y, en cierta manera, enigmáticas. Copérnico afirma que estaba motivado por visiones divergentes sobre la precisión del sistema de Ptolomeo, por la visión de que la teoría de Ptolomeo era sólo una hipótesis conveniente y por el conflicto entre los seguidores de las teorías de Aristóteles y de Ptolomeo.

Copérnico conservó algunos principios de la astronomía de Ptolomeo. Utilizó la circunferencia como la curva básica sobre la cual iba a construirse su explicación de los movimientos de los cuerpos celestes. Más todavía, como Ptolomeo, utilizó el hecho de que el movimiento de los planetas debe construirse mediante una sucesión de movimientos con velocidad constante. Su razón era que un cambio en la

velocidad podía ser motivado sólo por un cambio en la potencia motora y puesto que Dios, la causa del movimiento, era constante, el efecto lo tenía que ser también. Asumió el esquema griego del movimiento epicíclico sobre un deferente. Pero Copérnico rechazó el movimiento ecuante uniforme utilizado por Ptolomeo porque este movimiento no requiere una velocidad lineal uniforme.

Utilizando la idea de Aristarco de situar el Sol en lugar de la Tierra en el centro de cada deferente, Copérnico fue capaz de sustituir los complicados diagramas que se requerían anteriormente para describir el movimiento de cada cuerpo celeste por diagramas mucho más simples. En lugar de 77 circunferencias, necesitó sólo 34 para explicar el movimiento de la Luna y de los seis planetas conocidos. Más adelante refinó algo este esquema situando al Sol cerca, pero no exactamente en el centro, del sistema.

La teoría de Copérnico no concordaba mejor con las observaciones que las modificaciones de entonces de la teoría de Ptolomeo. El mérito de su sistema fue más bien que hizo que el movimiento de la Tierra alrededor del Sol explicara las principales irregularidades del movimiento de los planetas sin utilizar tantos epiciclos. Además, su esquema trataba a todos los planetas de la misma forma en general, mientras que Ptolomeo había utilizado esquemas algo diferentes para los planetas interiores, Mercurio y Venus, que para los exteriores, Marte, Júpiter y Saturno. Por último, el cálculo de las posiciones de los cuerpos celestes era más sencillo en el esquema de Copérnico, tanto que incluso en 1542 algunos astrónomos, utilizando su teoría, comenzaron la preparación de nuevas tablas de las posiciones celestes.

La teoría de Copérnico encontró una oposición tan profunda como llena de prejuicios. Las discrepancias entre la teoría de Copérnico y las observaciones hizo que Tycho Brahe (1546-1601) abandonara la teoría y buscara un compromiso. Vieta, por la misma razón, la rechazó también y volvió a intentar mejorar la teoría de Ptolomeo. La mayor parte de los intelectuales rechazaron la teoría, bien porque no llegaron a entenderla o porque no apoyaban ideas revolucionarias. Las matemáticas de la misma eran ciertamente difíciles de entender; como dice el mismo Copérnico en el prefacio, el libro estaba dirigido a matemáticos. La observación de una nueva estrella por Brahe y astrónomos alemanes en 1572 ayudó algo a la teoría. La súbita aparición y desaparición de estrellas contradecía el dogma aristotélico y escolástico de la invariabilidad de los cielos.

El destino de la teoría heliocéntrica hubiera sido muy incierto si

no es por el trabajo de Johannes Kepler (1571-1630). Nació en Weil, una ciudad en el ducado de Würtemberg. Su padre era un borrachín que pasaba de ser un mercenario a atender una taberna. Se le sacó pronto de la escuela y fue enviado a trabajar al campo. Cuando era todavía un niño, contrajo la viruela, que le dejó las manos lisiadas y la vista deteriorada. Sin embargo, se las arregló para obtener un grado universitario en el Colegio de Maulbronn en 1588; a continuación, dirigido hacia el ministerio religioso, estudió en la Universidad de Tübingen. Allí, un cordial profesor de matemáticas y de astronomía, Michael Mästlin (Möstlin, 1550-1631) le enseñó privadamente la teoría de Copérnico. Los superiores de Kepler en la universidad consideraron su dedicación y, en 1594, le ofrecieron un puesto de profesor de matemáticas y moral en la Universidad de Grätz, en Austria, que Kepler aceptó. Para realizar sus tareas se le exigió saber astrología; ello le volcó todavía más hacia la astronomía.

Kepler fue expulsado de Grätz cuando la ciudad pasó bajo control católico, y se convirtió en ayudante de Tycho Brahe en el observatorio de este último en Praga. Cuando murió, Kepler fue contratado en su lugar. Parte de su trabajo era elaborar horóscopos para quien le había contratado, el emperador Rodolfo II. Kepler se consolaba a sí mismo pensando que la astrología permitía vivir a los astrónomos.

Durante toda su vida Kepler estuvo acosado por todo tipo de dificultades. Su primera mujer y varios de sus hijos murieron. Como protestante, sufrió de diversas formas persecución por parte de los católicos. Estuvo con frecuencia en una situación económica desesperada. Su madre fue acusada de brujería y Kepler tuvo que defenderla. Sin embargo, a lo largo de todas estas desgracias continuó su trabajo científico con perseverancia, laboriosidad extraordinaria e imaginación fértil.

En su enfoque de los problemas científicos, Kepler es una figura de transición. Como Copérnico y los pensadores medievales, estaba atraído por una teoría bella y racional. Aceptó la doctrina platónica de que el universo está ordenado de acuerdo con un plan matemático preestablecido. Pero, a diferencia de sus predecesores, tenía un gran respeto por los hechos. Su trabajo más maduro estuvo basado enteramente en hechos, y en él avanzó desde los hechos a las leyes. En la búsqueda de leyes mostró una gran inventiva en las hipótesis, un amor por la verdad y una viva fantasía que no obstruía la razón. Aunque imaginó un gran número de hipótesis, no dudó en rechazarlas cuando no se adaptaban a los hechos.

Motivado por la belleza y armonía del sistema de Copérnico, decidió dedicarse a la búsqueda de las armonías geométricas adicionales que pudieran permitir explicar las observaciones mucho más precisas que había proporcionado Tycho Brahe. Su búsqueda de las relaciones matemáticas de cuya existencia estaba convencido le condujo a emplear años de trabajo siguiendo caminos falsos. En el prefacio de su *Mysterium Cosmographicum* (1596), dice: «Intento demostrar que Dios, al crear el universo y regular el orden del cosmos, tenía en su mente los cinco cuerpos regulares de la geometría tal como se conocen desde los tiempos de Pitágoras y Platón, y que fijó, de acuerdo con aquellas dimensiones, el número de cielos, sus proporciones y las relaciones de sus movimientos.»

Y así postuló que los radios de las órbitas de los seis planetas eran los radios de las esferas relacionadas con los cinco sólidos regulares de la siguiente manera. El radio más grande era el de la órbita de Saturno. En una esfera de este radio suponía inscrito un cubo. En este cubo se inscribía una esfera cuyo radio sería el de la órbita de Júpiter. En esta esfera suponía inscrito un tetraedro y en éste, a su vez, otra esfera, cuyo radio era el de la órbita de Marte, y así sucesivamente con los cinco sólidos regulares. Ello permitía seis esferas, justo lo suficiente para el número de planetas conocidos entonces. Sin embargo, las deducciones que se obtenían de esta hipótesis no estaban de acuerdo con las observaciones y abandonó la idea, pero no antes de que hubiera hecho extraordinarios esfuerzos para aplicarla, incluso en formas modificadas.

Aunque el intento de utilizar los cinco sólidos regulares para descubrir los secretos de la naturaleza no tuvo éxito, Kepler obtuvo un completo triunfo en sus esfuerzos posteriores para encontrar relaciones matemáticas armoniosas. Sus resultados más famosos e importantes se conocen hoy como las tres leyes de Kepler del movimiento planetario. Las dos primeras fueron publicadas en un libro de 1609 con un título muy largo que a veces se reduce a *Astronomia Nova* y a veces a *Comentarios sobre el movimiento de Marte*.

La primera ley establece que la trayectoria de cada planeta no es la resultante de la combinación de circunferencias que se mueven, sino que es una elipse con el Sol en uno de sus focos (fig. 12.2). La segunda ley de Kepler se entiende mejor mediante un diagrama (fig. 12.3). Los griegos, según vimos, creían que el movimiento de un planeta debe explicarse en términos de velocidades lineales constantes. Kepler,

como Copérnico, se aferró al principio a la doctrina de las velocidades constantes. Pero sus observaciones le impulsaron a abandonar también esta preciada creencia. Su alegría fue grande cuando fue capaz de sustituirla por algo igualmente atractivo, porque así se confirmaba su convicción de que la naturaleza sigue leyes matemáticas. Si *MM'* y *NN'* son distancias recorridas por un planeta en intervalos iguales de tiempo, entonces, de acuerdo con el principio de velocidad constante, *MM'* y *NN'* tendrían que ser distancias iguales. Sin embargo, de acuerdo con la segunda ley de Kepler, *MM'* y *NN'* no son iguales en general, pero las *áreas SMM'* y *SNN'* son iguales. Por tanto Kepler sustituyó distancias iguales por áreas iguales. Arrancar tal secreto a los planetas era un triunfo, porque la relación mencionada no se deduce en absoluto tan fácilmente como puede parecer una vez que se describe en un papel.

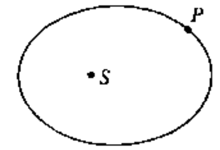


FIGURA 12.2

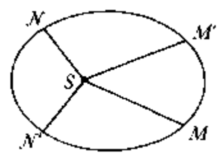


FIGURA 12.3

Kepler realizó esfuerzos todavía más extraordinarios para obtener la tercera ley del movimiento. Esta ley dice que el cuadrado del período de revolución de cualquier planeta es igual al cubo de su distancia media del Sol, con tal que se tomen como unidades de tiempo y de distancia [4] el período de revolución de la Tierra y su distancia al Sol. Kepler publicó este resultado en *La armonía del mundo* (1619).

El trabajo de Kepler es bastante más revolucionario que el de Copérnico; igualmente desafiante al adoptar el heliocentrismo, Kepler rompió radicalmente con la autoridad y la tradición al utilizar la elipse (oponiéndose a la composición de movimientos circulares) y velocidades no uniformes. Se mantuvo firmemente en la posición de que las investigaciones científicas son independientes de todas las doctrinas filosóficas y teológicas, que sólo las consideraciones mate-

[4] Aunque Kepler la estableció de esta manera, la formulación correcta requiere sustituir distancia media por semieje mayor.

máticas deben determinar la validez de cualquier hipótesis y que las hipótesis y la deducciones que se obtengan de ellas deben de resistir la confirmación empírica.

El trabajo de Copérnico y Kepler es notable en muchos aspectos, pero vamos a limitarnos a su importancia en la historia de las matemáticas. En vista de los muchos y serios argumentos en contra de la teoría heliocéntrica, su trabajo demostró cuán fuertemente había arraigado en Europa el punto de vista griego de que las verdades de la naturaleza reposan en leyes matemáticas.

Había objeciones científicas de peso, muchas de las cuales habían sido ya avanzadas por Ptolomeo, en contra de la sugerencia de Aristarco. ¿Cómo podía un cuerpo tan pesado como la Tierra ser puesto y mantenido en movimiento? Los otros planetas estaban en movimiento, incluso de acuerdo con la teoría de Ptolomeo, pero los griegos y pensadores medievales habían mantenido que éstos estaban compuestos de alguna sustancia ligera especial. Había otras objeciones. ¿Por qué, si la Tierra gira de Oeste a Este, al tirar un objeto al aire éste no cae al oeste de su posición original? ¿Por qué la Tierra no se deshace en su rotación? La respuesta, muy endeble, de Copérnico a la última objeción fue que la esfera es una forma natural y que se mueve naturalmente, y que por tanto la Tierra no se destruiría a sí misma. Se le preguntó también por qué los objetos que están sobre la Tierra, incluso el mismo aire, permanecen en ella si ésta gira aproximadamente a 3/10 de milla por segundo y se traslada alrededor del Sol a una velocidad aproximada de 18 millas por segundo. Si, como creían Ptolomeo y Copérnico, la velocidad de un cuerpo en movimiento natural es proporcional a su peso, la Tierra debía dejar detrás de sí objetos de peso menor. Copérnico replicó que el aire tenía «tendencia a la Tierra» y que, por lo tanto, permanecería con ella.

Todavía había objeciones científicas adicionales por parte de los astrónomos. Si la Tierra se movía, ¿por qué no cambiaba la dirección de las estrellas «fijas»? Un ángulo de paralaje de 2′ requería que la distancia de las estrellas fuera al menos cuatro millones de veces el radio de la Tierra; tal distancia era inconcebible en esa época. Al no detectar ningún paralaje de las estrellas (lo que implicaba que tenían que estar incluso más lejos), Copérnico declaró que «los cielos son inmensos comparados con la Tierra y parecen ser de tamaño infinito... Los límites del universo son desconocidos e imposibles de conocer». Entonces, dándose cuenta de lo inadecuado de su respuesta, propuso el problema a los filósofos y se evadió de él. Nadie midió este paralaje

hasta que en 1838 el matemático Bessel midió el correspondiente a una de las estrellas más próximas, y obtuvo que era de 0,31".

Si Copérnico y Kepler hubieran sido hombres «sensibles», no hubieran desafiado nunca a sus sentidos. No sentimos ni la rotación ni la revolución de la Tierra, a pesar de las grandes velocidades a que se realizan. Por otra parte, lo que sí vemos es el movimiento del Sol.

Copérnico y Kepler eran muy religiosos; sin embargo, ambos negaban una de las doctrinas centrales del cristianismo, cual es que el hombre, la principal preocupación de Dios, estaba en el centro del universo, y que todo en el universo giraba en torno de él. En contraste, la teoría heliocéntrica, al poner al Sol en el centro del universo, minaba este reconfortante dogma de la Iglesia, porque hacía aparecer al hombre como sólo uno de una multitud de peregrinos a la deriva a través de un frío firmamento. Parecía menos probable que hubiera nacido para vivir gloriosamente y alcanzar el paraíso después de su muerte. Menos probable, también, era que fuera objeto de los cuidados de Dios. Por tanto, al desplazar la Tierra, Copérnico y Kepler removieron una piedra angular de la teología católica y pusieron en peligro su estructura. Copérnico señaló que el universo es tan inmenso comparado con la Tierra que hablar de un centro no tiene sentido. Sin embargo, este razonamiento le colocó todavía más en oposición con la religión.

En contra de todas estas objeciones, Copérnico y Kepler tenían sólo una respuesta, pero de peso. Cada uno había obtenido una simplificación matemática y, en verdad, una teoría estéticamente superior y abrumadoramente armoniosa. Si las relaciones matemáticas eran el objetivo del trabajo científico, y si podía darse una descripción matemática mejor, entonces este hecho, reforzado con la creencia de que Dios había diseñado el mundo y habría utilizado, como es lógico, la teoría superior, bastaría para contrapesar todas las objeciones. Cada uno de ellos sintió, y estableció con claridad, que su trabajo revelaba la armonía, simetría y designio del taller divino y la todopoderosa presencia de Dios. Copérnico no podía contener su júbilo: «Obtenemos, por tanto, bajo esta disposición ordenada, una maravillosa simetría en el universo y una relación definida de armonía en el movimiento y magnitud del orbe, de una clase que no es posible obtener de otra manera.» El mismo título del trabajo de Kepler de 1619, *La armonía del mundo,* y las interminables alabanzas a Dios, expresando su satisfacción por la magnificencia del plan matemático de Dios, atestiguan sobre sus creencias.

No es sorprendente que, al principio, sólo matemáticos apoyaran la teoría heliocéntrica. Sólo un matemático, y uno convencido de que el universo estaba trazado matemáticamente, habría tenido la fortaleza mental para desdeñar las creencias filosóficas, religiosas y físicas que prevalecían entonces. Hasta que Galileo enfocó su telescopio hacia el firmamento, la evidencia astronómica no apoyó al razonamiento matemático. Las observaciones de Galileo, realizadas a principios del siglo XVII, revelaron cuatro lunas alrededor de Júpiter, mostrando que los planetas podían tener lunas. Se deducía, por tanto, que la Tierra podía no ser más que un planeta, precisamente porque tenía una luna. Galileo también observó que la Luna tenía una superficie rugosa y montañas y valles como la Tierra. En consecuencia, era también posible que la Tierra fuera sólo un cuerpo celeste más, y no necesariamente el centro del universo.

La teoría heliocéntrica ganó aceptación finalmente porque era más simple para los cálculos, por su superioridad matemática y porque la apoyaban las observaciones. Esto significó que la ciencia del movimiento tenía que rehacerse a la luz de una Tierra con movimiento de rotación y de revolución. En resumen, se necesitaba una nueva ciencia de la mecánica.

Las investigaciones sobre la luz y la óptica continuaron en una línea sin fisuras con respecto a lo que ya hemos señalado sobre el período medieval. En el siglo XVI los astrónomos se interesaron más por estos temas porque el efecto de refracción del aire sobre la luz cambia la dirección de los rayos de luz cuando éstos vienen de los planetas y de las estrellas y, por tanto, proporcionan una información equivocada sobre las direcciones de esos cuerpos. Hacia el final del siglo XVI fueron inventados el telescopio y el microscopio. Los usos científicos de esos instrumentos son obvios; abrieron nuevos mundos, y el interés por la óptica, ya extenso, se intensificó todavía más. Casi todos los matemáticos del siglo XVII realizaron trabajos sobre la luz y las lentes.

6. Notas sobre el Renacimiento

El Renacimiento no produjo ningún nuevo resultado brillante en matemáticas. Los pequeños progresos en este área contrastan con las realizaciones en literatura, pintura y arquitectura, en las que fueron creadas obras maestras que todavía forman parte de nuestra cultura, y

en la ciencia, en la que la teoría heliocéntrica eclipsó lo mejor de la astronomía griega y empequeñeció cualquier contribución árabe o medieval. Para las matemáticas este período fue, sobre todo, de absorción de los trabajos griegos. No fue tanto un renacimiento como una recuperación de una cultura más antigua.

Igualmente importante para la salud y crecimiento de las matemáticas fue que restableció, una vez más, como en los tiempos de Alejandría, sus conexiones íntimas con la ciencia y la tecnología. En la ciencia, el darse cuenta de que el objetivo eran las leyes matemáticas, el serlo todo y el fin de todo y, en tecnología, la valoración de que la formulación matemática de los resultados de las investigaciones era la más profunda y útil forma de conocimiento y la guía más segura para el diseño y la construcción, garantizaron el que las matemáticas fueran a ser una fuerza importante en los tiempos modernos y sentaron la promesa de nuevos desarrollos.

Bibliografía

Armitage, Angus: *Copernicus*, W. W. Norton, 1938.

—: *John Kepler*, Faber and Faber, 1966.

—: *Sun, Stand Then Still*, Henry Schumann, 1937; en rústica como *The World of Copernicus*, New American Library, 1951.

Ball, W. W. Rouse: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, cap. 12.

Baumgardt, Carola: *Johannes Kepler, Life and Letters*, Victor Gollancz, 1952.

Berry, Arthur: *A Short History of Astronomy*, Dover (reimpresión), 1961, pp. 86-197.

Braunmühl, A. von: *Vorlesungen über die Geschichte der Trigonometrie*, 2 vols., B. G. Teubner, 1900 y 1903, reimpreso por M. Sändig, 1970.

Burtt, E. A.: *Los fundamentos metafísicos de la Ciencia Moderna*, Buenos Aires, Suramericana, 1945.

Butterfield, Herbert: *Los Orígenes de la Ciencia Moderna*, Madrid, Taurus, 1982.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1900, vol. 2, pp. 1-344.

Caspar, Max: *Johannes Kepler*, trad. Doris Hellman, Abelard-Schumann, 1960.

Cohen, I. Bernard: *El nacimiento de una nueva física*, Madrid, Alianza, 1989.

Coolidge, Julian L.: *The Mathematics of Great Amateurs*, Dover (reimpresión), 1963, caps. 3-5.

Copernicus, Nicolaus: *Sobre las revoluciones de los orbes celestes*, Madrid, Tecnos, 1987.
Crombie, A. C.: *Historia de la ciencia de Agustín a Galileo*, Madrid, Alianza, 1979.
Da Vinci, Leonardo: *Philosophical Diary*, Philosophical Library, 1959.
—: *Tratado de la pintura*, Madrid, Akal, 1986.
Dampier-Whetham, W. C. D.: *Historia de la Ciencia*, Madrid, Tecnos, 1986.
Dijksterhuis, E. J.: *The Mechanization of the World Picture*, Oxford University Press, 1961, partes 3 y 4.
Drake, Stillman y I. E. Drabkin: *Mechanics in Sixteen-Century Italy*, University of Wisconsin Press, 1969.
Dreyer, J. L. E.: *A History of Astronomy from Thales to Kepler*, Dover (reimpresión), 1953, caps. 12-16.
—: *Tycho Brahe, A Picture of Scientific Life and Work in the Sixteen Century*, Dover (reimpresión), 1963.
Gade, John A.: *The Life and Times of Tycho Brahe*, Princeton University Press, 1947.
Galilei, Galileo: *Diálogo sobre los máximos sistemas*, Madrid, Alianza (en preparación).
Hall, A. R.: *La Revolución científica*, Barcelona, Crítica, 1985.
Hallerberg, Arthur E.: «George Mohr and *Euclidis Curiosi*», *The Mathematics Teacher*, 53, 1960, 127-132.
—: «The Geometry of the Fixed Compass», *The Mathematics Teacher*, 52, 1959, 230-244.
Hart, Ivor B.: *The World of Leonardo da Vinci*, Viking Press, 1962.
Hofmann, Joseph E.: *The History of Mathematics*, Philosophical Library, 1957.
Hughes, Barnabas: *Regiomontanus on Triangles*, University of Wisconsin Press, 1967. Es una traducción del *De Triangulis*.
Ivins, W. M., Jr.: *Art and Geometry* (1946). Dover (reimpresión), 1965.
Kepler, Johannes: *Gesammelte Werke*, C. H. Beck'sche Verlagsbuchhandlung, 1938-59.
—: *Concerning the More Certain Foundations of Astrology*, Cancy Publications, 1942. Muchos de los libros de Kepler han sido reimpresos y sólo pocos traducidos.
Koyré, Alexandre: *Del mundo cerrado al universo infinito*, Madrid, Siglo XXI, 1984.
—: *La Révolution astronomique*, Hermann, 1961.
Kuhn, Thomas S.: *La revolución copernicana*, Ariel, 1985.
MacCurdy, Edward: *The Notebooks of Leonardo da Vinci*, George Braziller (reimpresión), 1954.
Pannekoek, A.: *A History of Astronomy*, John Wiley and Sons, 1961, caps. 16-25.

Panofsky, Erwin: *«Dürer as a Mathematician»*, en James R. Newman, *The World of Mathematics,* Simon and Schuster, 1956, pp. 603-621.

Santillana, G. de: *The Crime of Galileo*, University of Chicago Press, 1955.

Sarton, George: *The Appreciation of Ancient and Medieval Science During the Renaissance*, University of Pennsylvania Press, 1957.

—: *Six Wings: Men of Science in the Renaissance*, Indiana University Press, 1957.

Smith, David Eugene: *History of Mathematics*, Dover (reimpresión), 1958, vol. I, cap. 8, y vol. 2, cap. 8.

Smith, Preserved: *A History of Modern Culture,* Holt, Rinehart and Winston, 1940, vol. I, caps. 2-3.

Taylor, Henry Osborn: *Thought and Expression in the Sixteenth Century,* Crowell-Collier (reimpresión), 1962, parte V.

Tropfke, Johannes: *Geschichte der Elementarmathematik*, 7 vols., 2.ª ed., W. de Gruyter, 1921-24.

Vasari, Giorgio: *Vidas de artistas ilustres*, Madrid, Iberia, 1957.

Wolf, Abraham: *A History of Science, Technology and Philosophy in the Sixteenth and Seventeenth Centuries,* George Allen and Unwin, 1950, caps. 1-6.

Zeller, Sister Mary Claudia: «The Development of Trigonometry from Regiomontanus to Pitiscus», Ph. D. Dissertation (Tesis Doctoral), University of Michigan, 1944; Edwards Brothers, 1946.

Capítulo 13
LA ARITMETICA Y EL ALGEBRA EN LOS SIGLOS XVI Y XVII

El álgebra es el instrumento intelectual que aclara los aspectos cuantitativos del mundo.

ALFRED NORTH WHITEHEAD

1. Introducción

Los primeros desarrollos matemáticos europeos de importancia tuvieron lugar en la aritmética y el álgebra. El trabajo de hindúes y árabes había puesto los cálculos aritméticos prácticos en primera línea matemática y había fundado el álgebra sobre una base aritmética en vez de geométrica. Este trabajo atrajo también la atención sobre el problema de la resolución de ecuaciones.

En la primera mitad del siglo XVI apenas hubo cambio alguno con respecto a la actitud de espíritu de los árabes, sino solamente un incremento en el tipo de actividad que los europeos habían aprendido de las obras árabes. Para mediados de siglo, las necesidades prácticas y científicas de la civilización europea exigían ya más avances en aritmética y álgebra. Las aplicaciones tecnológicas del trabajo científico y las necesidades prácticas requerían, como ya hemos apuntado, resultados cuantitativos. Por ejemplo, las lejanas exploraciones geográficas precisaban un conocimiento astronómico más exacto. Al mismo tiempo, el interés en conectar la nueva teoría astronómica con las cada vez más precisas observaciones exigía mejores tablas astronómicas, lo que, a su vez, significaba disponer de tablas trigonométricas

más precisas. El desarrollo de la actividad bancaria y comercial pedía una mejor aritmética. La respuesta a estos intereses es evidente en los escritos de Pacioli, Tartaglia y Stevin, entre otros. La *Summa* de Pacioli y el *General trattato de' numeri e misure* (1556) de Tartaglia contienen un número inmenso de problemas de aritmética mercantil. Finalmente, el trabajo técnico de los artesanos, especialmente en arquitectura, la fabricación de cañones y el movimiento de proyectiles necesitaban un nuevo pensamiento cuantitativo. Además de estas aplicaciones, una utilización totalmente nueva del álgebra, la representación de curvas, motivó enormes cantidades de trabajo. Bajo la presión de estas necesidades se aceleró el progreso en álgebra.

Vamos a dividir el análisis de todos estos nuevos desarrollos en cuatro epígrafes: aritmética, simbolismo, teoría de ecuaciones y teoría de números.

2. La situación del sistema numérico y la aritmética

Hacia el año 1500 se aceptaba el cero como un número y los números irracionales se usaban con más libertad. Pacioli, el matemático alemán Michael Stifel (1486?-1567), el ingeniero militar Simon Stevin (1548-1620) y Cardano utilizaban números irracionales en la tradición de hindúes y árabes, introduciendo cada vez más tipos. Stifel, por ejemplo, trabajaba con irracionales de la forma $\sqrt[m]{a + \sqrt[n]{b}}$, y Cardano racionalizaba fracciones con raíces cúbicas. La medida en que se llegaron a utilizar los números irracionales viene ejemplificada por la expresión de Vieta para π [1]. Considerando polígonos regulares de 4, 8, 16... lados inscritos en un círculo de radio unidad, Vieta halló que el valor de π viene dado por la expresión

$$\frac{2}{\pi} = \cos\frac{90^\circ}{2}\cos\frac{90^\circ}{4}\cos\frac{90^\circ}{8}\ldots$$

$$= \sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}}\ldots$$

Aunque los cálculos con números irracionales se efectuaban con

[1] El símbolo π fue usado por primera vez por William Jones (1706).

libertad, el problema de si tales expresiones eran realmente números era aún fuente de inquietud. En su obra principal, la *Arithmetica Integra* (1544), que trata de aritmética, de los irracionales del libro X de Euclides y de álgebra, Stifel considera la expresabilidad de los irracionales en notación decimal. Por una parte, argumenta:

> Dado que, al analizar figuras geométricas, cuando nos fallan los números racionales toman su lugar los irracionales y prueban exactamente las cosas que los números racionales no pudieron probar... nos vemos movidos y obligados a afirmar que son verdaderamente números; obligados, esto es, por los resultados que se siguen de su uso, resultados que percibimos como reales, ciertos y constantes. Por otra parte, otras consideraciones nos obligan a negar que los números irracionales sean números en absoluto. Esto es, cuando tratamos de someterlos a numeración [representación decimal]... hallamos que se escapan continuamente, de forma que ninguno de ellos puede ser aprehendido precisamente en sí mismo... Y nada de tal naturaleza carente de precisión puede llamarse número... Por consiguiente, de la misma forma que un número infinito no es un número, un número irracional no es un número verdadero, sino que yace oculto en una especie de nube de infinitud.

A continuación, argumenta que los números son enteros o fraccionarios; obviamente, los irracionales no son ni una cosa ni otra, luego no son realmente números. Un siglo después, Pascal y Barrow decían que un número como $\sqrt{3}$ sólo puede entenderse como una magnitud geométrica; los números irracionales son meros símbolos que no tienen existencia independiente de la magnitud geométrica continua, y la lógica de las operaciones con números irracionales debe justificarse por el método eudoxiano de las magnitudes. Este era también el punto de vista de Newton en su *Arithmetica Universalis* (publicada en 1707, aunque basada en clases dadas treinta años antes).

Otros hicieron afirmaciones positivas de que los números irracionales eran entidades independientes. Stevin consideraba los irracionales como números, y los aproximaba cada vez más por racionales; John Wallis, en su *Algebra* (1685), también aceptaba los irracionales como números en su pleno sentido. Consideraba el libro V de los *Elementos* de Euclides como de naturaleza esencialmente aritmética. También Descartes, en las *Reglas para la dirección del espíritu* (hacia 1628), admitía los irracionales como números abstractos que pueden representar magnitudes continuas.

En cuanto a los números negativos, aunque conocidos en Europa a través de los textos árabes, no eran aceptados como números por la

mayoría de los matemáticos de los siglos XVI y XVII, o, si lo eran, nunca como raíces de ecuaciones. Nicolás Chuquet (1445?-1500?) en el siglo XV y Stifel (1553) en el XVI hablaban de los números negativos como absurdos. Cardano daba números negativos como raíces de ecuaciones, pero los consideraba soluciones imposibles, meros símbolos, llamándolos «ficticios», mientras que a las raíces positivas las llamaba «reales». Vieta descartaba enteramente los números negativos. Descartes los aceptaba en parte: llamaba «falsas» a las raíces negativas de las ecuaciones, con el argumento de que pretendían representar números menores que la nada. Había, sin embargo, mostrado (ver la sección 5) que, dada una ecuación, es posible obtener otra cuyas raíces son mayores en una cantidad dada que las de la original, de forma que una ecuación con raíces negativas puede transformarse en otra con raíces positivas. Dado que podemos convertir raíces falsas en raíces reales, Descartes estaba dispuesto a aceptar los números negativos. Pascal consideraba sustraer 4 de 0 completamente absurdo.

Un interesante argumento en contra de los números negativos lo dio Antoine Arnauld (1612-94), teólogo y matemático, y buen amigo de Pascal. Arnauld cuestionaba que $-1 : 1 = 1 : -1$, ya que, según decía, -1 es menor que $+1$, y, por tanto, ¿cómo iba a ser un menor a un mayor como un mayor a un menor? Este problema fue discutido por muchos. En 1712, Leibniz concedía [2] que la objeción era válida, pero argumentaba que es posible calcular con tales proporciones pues su forma es correcta, de la misma forma que es posible calcular con cantidades imaginarias.

Uno de los primeros algebristas que aceptó los números negativos fue Thomas Harriot (1560-1621), que de vez en cuando ponía un número negativo sólo como segundo miembro de una ecuación, aunque no aceptaba raíces negativas. Rafael Bombelli (siglo XVI) dio definiciones claras para los números negativos. Stevin utilizaba coeficientes positivos y negativos en las ecuaciones, y aceptaba también raíces negativas. En su *L'Invention nouvelle en l'algèbre* (1629), Albert Girard (1595-1632) colocaba los números negativos en paridad con los positivos, y daba las dos raíces de la ecuación de segundo grado, incluso si ambas eran negativas. Tanto Girard como Harriot usaban el signo «menos» para la operación de sustracción y para los números negativos.

---

[2] *Acta Erud.*, 1712, 167-69 = *Math. Schriften*, 5, 387-89.

En términos globales, no muchos matemáticos de los siglos XVI y XVII se sentían a gusto o aceptaban los números negativos como tales, por no hablar de reconocerlos como verdaderas raíces de las ecuaciones. Había ciertas creencias curiosas acerca de ellos. Aunque Wallis se adelantó a su tiempo aceptándolos, creía que eran mayores que infinito, pero no menores que cero. En su *Arithmetica Infinitorum* (1655), razonaba que, dado que la razón $a/0$, con a positivo, es infinita, al cambiar el denominador por una cantidad negativa, como en a/b con b negativo, la razón debe ser mayor que infinito.

Aun sin haber vencido completamente sus dificultades con los números irracionales y negativos, los europeos se buscaron más problemas al toparse con lo que hoy llamamos números complejos. Obtuvieron dichos números extendiendo la operación aritmética de la raíz cuadrada a números cualesquiera que apareciesen al resolver ecuaciones de segundo grado por el método usual de completar el cuadrado. Así, Cardano, en el capítulo 37 de su *Ars Magna* (1545), plantea y resuelve el problema de dividir 10 en dos partes cuyo producto sea 40, cuya ecuación es $x(10 - x) = 40$. Obtiene las raíces $5 + \sqrt{-15}$ y $5 - \sqrt{-15}$, y luego dice: «dejando a un lado las torturas mentales que ello implica», multipliquemos $5 + \sqrt{-15}$ y $5 - \sqrt{-15}$; el producto es $25 - (-15)$, es decir, 40. Afirma entonces que «así progresa la sutileza aritmética, cuyo fin es, como se ha dicho, tan refinado como inútil». Como pronto veremos, Cardano se vio aún más comprometido con los números complejos al resolver la ecuación de tercer grado (sección 4). También Bombelli consideraba números complejos en la solución de ecuaciones de tercer grado y formuló en forma prácticamente moderna las cuatro operaciones con números complejos, aunque todavía los consideraba como inútiles y «sofísticos». Albert Girard sí reconocía los números complejos, al menos como soluciones formales de ecuaciones. En *L'Invention nouvelle en l'algèbre*, dice: «Se puede decir: ¿Por qué son útiles esas soluciones imposibles [raíces complejas]? A lo que yo respondo: por tres cosas: por la certeza de las reglas generales, por su utilidad, y porque no hay otras soluciones.» Sin embargo, los avanzados puntos de vista de Girard no tuvieron influencia.

También Descartes rechazó las raíces complejas, acuñando para ellas el término «imaginarias». En *La Géometrie* dice: «Ni las raíces verdaderas ni las falsas [negativas] son siempre reales; a veces son imaginarias.» Razonaba que, mientras que las raíces negativas al menos pueden hacerse «reales» transformando la ecuación en la que

aparecen en otras cuyas raíces sean positivas, esto no puede hacerse para las raíces complejas. Estas, por tanto, no son reales sino imaginarias; no son números. Descartes hacía una distinción más clara que sus predecesores entre raíces reales e imaginarias de las ecuaciones.

El propio Newton no consideraba las raíces complejas como significativas, probablemente porque en su tiempo carecían de sentido físico. Dice en *Universal Arithmetic* [3]: «Es de razón que las raíces de las ecuaciones sean imposibles [complejas], no vaya a ser que presenten casos de problemas que son imposibles como si fuesen posibles.» Es decir, los problemas que no tienen solución física o geométricamente real deberían tener raíces complejas.

Esta falta de claridad acerca de los números complejos puede ilustrarse por el muy citado aserto de Leibniz: «El Espíritu Divino halló una sublime expresión en esa maravilla del análisis, ese portento del mundo ideal, ese anfibio entre el ser y el no ser que llamamos la raíz imaginaria de la unidad negativa» [4]. Aunque Leibniz trabajaba formalmente con números complejos, no entendía su naturaleza.

Durante los siglos XVI y XVII, los métodos operativos con números reales fueron mejorados y extendidos. En Bélgica (a la sazón parte de los Países Bajos) encontramos a Stevin en *La Disme* (Aritmética Decimal, 1585) abogando por el uso de los decimales, en oposición al sistema sexagesimal, para escribir y operar con fracciones. Otros, como Christoff Rudolff (*c.* 1550-*c.* 1545), Vieta y el árabe al-Kashî (m. hacia 1436), los habían utilizado previamente. Stevin recomendaba la adopción de un sistema decimal de pesos y medidas, en la idea de ahorrar tiempo y trabajo a los contables (él mismo había comenzado su carrera como escribiente). Escribía 5,912 como 5 ⓪ 9 ① 1 ② 2 ③, o como 5, 9′ 1″ 2‴. Vieta mejoró y extendió los métodos de extracción de raíces cuadradas y cúbicas.

El uso de las fracciones continuas en la aritmética constituye otro de los desarrollos de este período. Recordemos que los hindúes, en particular Āryabhata, habían utilizado fracciones continuas para resolver ecuaciones lineales indeterminadas. Bombelli, en su *Algebra* (1572), fue el primero en usarlas para aproximar raíces cuadradas. Para aproximar $\sqrt{2}$ escribe

[3] Segunda edición, 1728, p. 193.

[4] *Acta Erud.*, 1702 = *Math. Schriften*, 5, 350-61.

$$\sqrt{2} = 1 + \frac{1}{y} \tag{1}$$

obteniendo de aquí

$$y = 1 + \sqrt{2}. \tag{2}$$

Sumando 1 a ambos miembros de (1) y usando (2) se tiene

$$y = 2 + \frac{1}{y} \cdot \tag{3}$$

Por tanto, de nuevo por (1) y (3)

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{y}}.$$

Y, como y está dado por (3)

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{y}}}.$$

Por sustitución repetida del valor de y, Bombelli obtiene

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2} + \ldots}}}$$

El segundo miembro se escribe también de la forma

$$\sqrt{2} = 1 + \frac{1}{2+}\,\frac{1}{2+}\,\frac{1}{2+}\ldots$$

Esta fracción continua es simple porque todos los numeradores son 1, y es periódica porque los denominadores se repiten. Bombelli dio otros ejemplos de cómo obtener fracciones continuas, pero no consideró la cuestión de si los desarrollos convergían hacia los números que se suponía representaban.

El matemático inglés John Wallis, en su *Arithmetica Infinitorum* (1655), representa $\frac{4}{\pi}$ como el producto infinito $\frac{3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdots}{2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots}$. En este libro afirma igualmente que Lord William Brouncker (1620-84), primer presidente de la Real Sociedad (Royal Society), había transformado este producto en la fracción continua

$$\frac{4}{\pi} = 1 + \frac{1}{2+}\,\frac{9}{2+}\,\frac{25}{2+}\,\frac{49}{2+} \cdots$$

Brouncker no volvió a utilizar esta forma, pero Wallis prosiguió el trabajo. En su *Opera Mathematica*, I (1695), en la que introdujo el término «fracción continua», dio la regla general para calcular las convergentes de una fracción continua. Si p_n/q_n es la convergente n-sima de la fracción continua

$$\frac{b_1}{a_1+}\,\frac{b_2}{a_2+}\,\frac{b_3}{a_3+} \ldots, \quad \text{entonces} \quad \frac{p_n}{q_n} = \frac{a_n p_{n-1} + b_n p_{n-2}}{a_n q_{n-1} + b_n q_{n-2}}.$$

En esta época no se obtuvo ningún resultado definitivo con respecto a la convergencia de p_n/q_n hacia el número representado por la fracción continua.

El mayor avance en la aritmética durante los siglos XVI y XVII fue la invención de los logaritmos. La idea básica fue indicada por Stifel. En *Arithmetica Integra*, observó que los términos de la progresión geométrica

$$1, r^2, r^3, \ldots$$

se corresponden como los términos de la progresión aritmética de los exponentes

$$0, 1, 2, 3, \ldots$$

La multiplicación de dos términos de la progresión geométrica da como resultado un término cuyo exponente es la suma de los correspondientes términos de la progresión aritmética, y la división de dos términos de la progresión geométrica da un término cuyo exponente es la diferencia de los correspondientes términos de la progresión aritmética. Esta observación la hace también Chuquet en *Le Triparty en la science des nombres* (1484). Stifel extendió esta relación entre las dos progresiones a los exponentes negativos y fraccionarios. Así, la división de r^2 por r^3 da r^{-1}, que corresponde al término -1 en la progresión aritmética extendida. Stifel, sin embargo, no utilizó esta conexión entre ambas progresiones para introducir los logaritmos.

John Napier (1550-1617), el escocés que desarrolló los logaritmos hacia 1594, fue guiado por esta correspondencia entre los términos de una progresión geométrica y los de la progresión aritmética correspondiente. A Napier le interesaba facilitar los cálculos de trigonometría esférica que precisaban los problemas astronómicos. De hecho, envió sus resultados preliminares a Tycho Brahe para su aprobación.

Napier explicó sus ideas en *Mirifici Logarithmorum Canonis Descriptio* (1614) y en *Mirifici Logarithmorum Canonis Constructio* (1619), publicada póstumamente. Como le interesaba la trigonometría esférica, consideró los logaritmos de senos. Siguiendo a Regiomontano, que utilizaba semicuerdas de un círculo cuyo radio contenía 10^7 unidades, Napier empezó con 10^7 como el mayor número considerado. Representó los valores del seno de 10^7 a 0 sobre la línea AZ (ver fig. 13.1) y supuso que A se mueve hacia Z con velocidad proporcional a su distancia de Z.

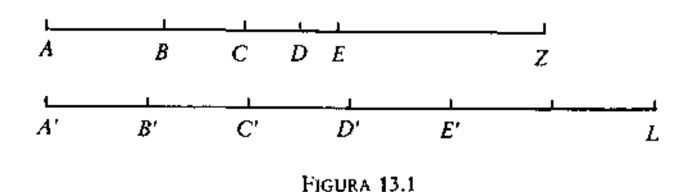


FIGURA 13.1

En sentido estricto, la velocidad del punto móvil varía continuamente con la distancia desde A, y su correcta expresión precisa del cálculo. Sin embargo, si consideramos un pequeño intervalo de tiempo t y hacemos que las longitudes AB, BC, CD, ... sean recorridas en ese intervalo, y suponemos que la velocidad durante el

intervalo t es constante e igual a la del punto móvil al principio del intervalo, entonces las longitudes AZ, BZ, CZ, ... están en progresión geométrica. En efecto, consideremos DZ, y sea k la constante de proporcionalidad que relaciona la velocidad y la distancia de Z del punto móvil. Se tiene

$$DZ = CZ - CD.$$

Además, la velocidad del punto móvil en C es $k(CZ)$. Entonces, la distancia CD es igual a $k(CZ)t$. Por tanto,

$$DZ = CZ - k(CZ)t = CZ(1 - kt).$$

Así pues, cada longitud de la sucesión AZ, BZ, CZ, ... es $1 - kt$ multiplicado por la distancia anterior.

A continuación, Napier considera otro punto que parte al mismo tiempo que A y se mueve a velocidad *constante* en la recta $A'L$ (fig. 13.1), extendida indefinidamente hacia la derecha, de forma que este punto alcanza B', C', D'... al mismo tiempo que el primer punto llega a B, C, D... respectivamente. Las distancias $A'B'$, $A'C'$, $A'D'$, ... están obviamente en progresión aritmética. Estas distancias $A'B'$, $A'C'$, $A'D'$, ... fueron consideradas por Napier como los logaritmos de BZ, CZ, DZ, ... respectivamente. El logaritmo de AZ ó 10^7 se tomaba como 0. Por tanto, los logaritmos crecen en progresión aritmética, mientras que los números (senos) decrecen en progresión geométrica. La cantidad aquí denotada como $1 - kt$ es $1 - \frac{1}{10^7}$ en la obra original de Napier, pero la cambió al llevar a cabo el cálculo de los logaritmos.

Obsérvese que cuanto menor tomemos t menor será el decrecimiento de los senos desde AZ hasta BZ, CZ, etc., y más próximos entre sí estarán los números de la tabla de logaritmos.

Napier tomó las distancias $A'B'$, $A'C'$ como 1, 2, 3, ... aunque no había necesidad de hacerlo. Podrían haber sido 1/2, 1, 1 1/2, 2, ... y el esquema habría funcionado exactamente igual. Además, los números originales eran significativos en cuanto que eran cantidades, con independencia de que fuesen senos, de forma que el esquema de Napier proporcionaba realmente logaritmos de números. El propio Napier aplicó los logaritmos a cálculos de trigonometría esférica.

La palabra «logaritmo», acuñada por Napier, significa «número de

la razón». «Razón» se refiere a la razón común de la sucesión de números *AZ*, *BZ*, *CZ*, ... También se refería a los logaritmos como «números artificiales».

Henry Briggs (1561-1631), profesor de matemáticas y astronomía, sugirió a Napier en 1615 que se utilizase 10 como base, y que el logaritmo de un número fuese el exponente en la potencia de 10 que igualase dicho número. Aquí, en oposición al esquema de Napier, se elige primero la base. Briggs calculó sus logaritmos tomando raíces cuadradas sucesivas de 10, es decir, $\sqrt{10}$, $\sqrt{\sqrt{10}}$, ..., hasta alcanzar, tras 54 extracciones de raíces, un número ligeramente mayor que 1. Es decir, obtuvo el número $A = 1010^{[(1/2)^{54}]}$. Tomó entonces $\log_{10}A$ como $(1/2)^{54}$. Usando el hecho de que el logaritmo de un producto de números es la suma de sus logaritmos, construyó una tabla de logaritmos de números muy próximos entre sí. Las tablas de logaritmos comunes actualmente en uso se derivan de las de Briggs.

Joost Bürgi (1552-1632), relojero e instrumentista suizo y ayudante de Kepler en Praga, se interesaba también en facilitar los cálculos astronómicos. Inventó los logaritmos, independientemente de Napier, hacia 1600, pero no publicó su trabajo, *Progress Tabulen*, hasta 1620. También Bürgi fue estimulado por las observaciones de Stifel de que la multiplicación y división de términos en una progresión geométrica puede llevarse a cabo sumando y restando los exponentes. Su obra aritmética fue similar a la de Napier.

Gradualmente fueron introduciéndose variantes de la idea de Napier. Se obtuvieron muchas y distintas tablas de logaritmos por medios algebraicos. El cálculo de logaritmos mediante series infinitas lo llevaron a cabo más tarde James Gregory, Lord Brouncker, Nicholas Mercator (n. Kaufman, 1620-87), Wallis y Edmond Halley (ver cap. 20, sec. 2).

Aunque la definición de los logaritmos como los exponentes de las potencias que representan un número en una base fija, según la idea de Briggs, se convirtió en el método usual, no se llegaron a definir como exponentes al comienzo del siglo XVII porque los exponentes fraccionarios e irracionales no se utilizaban. Hacia fin de siglo, una serie de matemáticos cayeron en la cuenta de que los logaritmos podían definirse de esa manera, pero la primera exposición sistemática de este enfoque no tuvo lugar hasta 1742, cuando William Jones (1675-1749) lo presentó en la introducción de la *Table of Logarithms* de William Gardiner. Euler había definido ya los logaritmos como exponentes, y en 1728, en un manuscrito no publicado (*Opera Posthuma*, II,

800-804), introdujo el número *e* como base de los logaritmos naturales.

El siguiente desarrollo en la aritmética (cuya realización ulterior ha resultado decisiva en tiempos recientes) fue la invención de instrumentos mecánicos y máquinas para acelerar la ejecución de los procesos aritméticos. La regla de cálculo procede del trabajo de Edmund Gunter (1581-1626), que utilizó los logaritmos de Napier. William Oughtred (1574-1660) introdujo reglas de cálculo circulares.

En 1642, Pascal inventó una máquina de calcular que hacía las sumas llevando de forma automática las cifras de las unidades a las decenas, de las decenas a las centenas, etc. Leibniz la vio en París e inventó a continuación una máquina de multiplicar. Mostró su idea a la Real Sociedad de Londres en 1677, y se publicó una descripción en la Academia de Berlín en 1710. A finales del siglo XVII, Samuel Morland (1625-95) inventó independientemente una máquina de sumar y restar, y otra de multiplicar.

No seguiremos con la historia de las máquinas de calcular porque hasta 1940, por lo menos, eran simplemente instrumentos mecánicos que llevaban a cabo operaciones aritméticas, y no tuvieron influencia en el curso de las matemáticas. Observemos, no obstante, que el paso más significativo entre las máquinas ya descritas y los modernos ordenadores electrónicos fue efectuado por Charles Babbage (1792-1871), que introdujo una máquina orientada al cálculo astronómico y de navegación. Su «máquina analítica» fue diseñada para llevar a cabo toda una serie de operaciones aritméticas basadas en instrucciones dadas a la máquina al comienzo, poniéndose luego a trabajar automáticamente mediante la fuerza del vapor. Con apoyo del gobierno británico, construyó modelos de demostración. Lamentablemente, la máquina planteaba demandas que eran excesivas para las posibilidades de la ingeniería de la época.

3. El simbolismo

Hubo un avance en el álgebra que resultó mucho más significativo para su propio desarrollo y el del análisis que el progreso técnico del siglo XVI, y fue la introducción de un mejor simbolismo. Ciertamente, este paso hizo posible hacer una ciencia del álgebra. Antes del siglo XVI, el único matemático que había introducido conscientemente el simbolismo para hacer más compacto y efectivo el razonamiento y la

escritura algebraica fue Diofanto. Todos los demás cambios de notación eran esencialmente abreviaturas de palabras normales introducidas de forma un tanto accidental. En el Renacimiento, el estilo habitual era aún retórico, con uso de palabras especiales, abreviaturas y, por supuesto, los símbolos de los números.

En el siglo XVI, la presión en pro de la introducción del simbolismo vino indudablemente de las crecientes demandas científicas que se ejercían sobre los matemáticos, de la misma forma que los avances en los métodos de cálculo lo fueron en respuesta al creciente uso de tales artes. Su progreso fue, sin embargo, intermitente. Muchos cambios se efectuaron por accidente, y es claro que los hombres del siglo XVI no llegaron a percibir lo que el simbolismo podría hacer en favor del álgebra. Ni siquiera los avances importantes en el simbolismo eran inmediatamente aceptados por los matemáticos.

Probablemente, las primeras abreviaturas, usadas del siglo XV en adelante, fueron p para «más» y m para «menos», pero en el Renacimiento, y especialmente en los siglos XVI y XVII, se introdujeron símbolos especiales. Los símbolos + y − fueron introducidos por los alemanes en el siglo XV para denotar excesos y defectos en los pesos de cofres y arcas, y los matemáticos los adoptaron, apareciendo en los manuscritos ya desde 1481. El símbolo × para «por» se debe a Oughtred, aunque Leibniz planteó la certera objeción de que podría confundirse con la letra x.

El siglo = fue introducido en 1557 por Robert Recorde (1510-58), de Cambridge, que escribió el primer tratado de álgebra, *The Whetstone of Witte* (1557). Decía que no conocía dos cosas más iguales que dos líneas paralelas, y por tanto dos líneas de ese tipo deberían denotar la igualdad. Vieta, que al principio escribía «aequalis», usó después ∼ para la igualdad. Descartes utilizaba ∝. Los símbolos > y < se deben a Thomas Harriot. Los paréntesis aparecen en 1544, y los corchetes y llaves, introducidos por Vieta, datan de 1593, aproximadamente. El símbolo de raíz cuadrada, $\sqrt{\ }$, era utilizado por Descartes, quien, sin embargo, escribía para la raíz cúbica $\sqrt{c}$.

Observemos unos ejemplos de formas de escritura. Usando R para la raíz cuadrada, p para «más» y m para «menos», Cardano escribía

$$(5+\sqrt{-15})\cdot(5-\sqrt{-15}) = 25-(-15) = 40$$

como

5p: Rm:15
5m: Rm:15
25m:m:15 qd. est 40.

También escribía

$\sqrt{7 + \sqrt{14}}$ como R.V. 7.p: R14.

La V indicaba que todo lo que seguía estaba bajo el signo radical.

El uso de símbolos para las incógnitas y sus potencias tuvo un ascenso sorprendentemente lento, si se tiene en cuenta la simplicidad y, sin embargo, extraordinario valor de tal práctica. (Naturalmente, Diofanto había usado tales símbolos.) Autores de comienzos del XVI, como Pacioli, se referían a la incógnita como *radix* («raíz» en latín) o *res* («cosa» en latín), *cosa* («cosa» en italiano) y *coss* («cosa» en alemán), razón por la cual el álgebra llegó a ser conocido como el arte «cósico». En su *Ars Magna,* Cardano se refería a la incógnita como *rem ignotam.* Escribía $x^2 = 4x + 32$ como *qdratu aeqtur* 4 *rebus p*:32 [5]. El término constante, 32, se llamaba el *número.* Los términos y notaciones variaban enormemente. Muchos símbolos se derivaban de abreviaturas: por ejemplo, un símbolo para la incógnita era *R*, abreviatura de *res*. La segunda potencia, representada por *Z* (de *zensus*) se llamaba *quadratum* o *censo.* *C*, tomado de *cubus*, denotaba x^3.

Los exponentes fueron gradualmente introducidos para denotar las potencias de *x*. Recordemos que ya Oresme, en el siglo XIV, usaba exponentes asociados a números. En 1484, Chuquet, en *Triparty,* escribía 12^3, 10^5 y 120^8 para indicar $12x^3$, $10x^5$ y $120x^8$. También escribía 12^0 por $12x^0$ y 7^{1m} por $7x^{-1}$. Así, 8^3, 7^{1m} igual a 56^2 significaba $8x^3 \cdot 7x^{-1} = 56x^2$.

En su *Algebra,* Bombelli utilizaba la palabra *tanto* en vez de *cosa.* Para designar x, x^2 y x^3 escribía $\underset{\smile}{1}$, $\underset{\smile}{2}$ y $\underset{\smile}{3}$. Así, $1 + 3x + 6x^2 + x^3$ es 1 *p.* $3^{\underset{\smile}{1}}$ *p.* $6^{\underset{\smile}{2}}$ *p.* $1^{\underset{\smile}{3}}$. En 1585, Stevin escribía esta expresión de la forma 1⓪ + 3① + 6② + ③. Stevin también usaba exponentes fraccionarios, $\frac{1}{2}$ para la raíz cuadrada, $\frac{1}{3}$ para la raíz cúbica, y así sucesivamente.

[5] *Rem* y *rebus* son formas correspondientes a la declinación de *res.*

Claude Bachet de Mézirac (1581-1638) prefería escribir $x^3 + 13x^2 + 5x + 2$ como $1C + 13Q + 5N + 2$. Vieta usaba la misma notación para las ecuaciones con coeficientes numéricos.

Descartes hizo un uso bastante sistemático de los exponentes enteros positivos. Expresaba

$$1 + 3x + 6x^2 + x^3 \quad \text{como} \quad 1 + 3x + 6xx + x^3.$$

Ocasionalmente, como otros, usaba también x^2. Para potencias superiores empleaba x^4, x^5, ... pero no x^n. Newton usaba exponentes positivos, negativos, enteros y fraccionarios, como $x^{5/3}$ y x^{-3}. Cuando en 1801 Gauss adoptó x^2 para xx, la primera de éstas se convirtió en la usual.

El cambio más significativo en el carácter del álgebra fue introducido por François Vieta en relación con el simbolismo. Educado como abogado, trabajó como tal en el parlamento de Bretaña. Más tarde fue consejero privado de Enrique de Navarra. Cuando, como resultado de problemas políticos, estuvo alejado de su cargo entre 1584 y 1589, se dedicó enteramente a las matemáticas. En general, se interesaba por ellas como entretenimiento, e imprimió e hizo circular su trabajo a sus expensas..., garantía de olvido, como dijo un escritor.

Vieta era un humanista en espíritu e intención; deseaba ser el conservador, redescubridor y continuador de la matemática antigua. Para él, innovación era renovación. Describe su *In Artem Analyticam Isagoge* [6] como «la obra del análisis matemático restaurado». Para escribir este libro se inspiró en la *Colección Matemática* de Pappus y en la *Arithmetica* de Diofanto. Creía que los antiguos habían empleado un tipo algebraico general de cálculo, que él reintrodujo en su álgebra, reactivando meramente así un arte conocido y aprobado en la antigüedad.

Durante el hiato en su carrera política, Vieta estudió las obras de Cardano, Tartaglia, Bombelli, Stevin y Diofanto. De ellas, y particularmente de la de Diofanto, extrajo la idea de emplear letras. Aunque una serie de matemáticos, incluyendo Euclides y Aristóteles, habían usado letras en lugar de números específicos, estos usos eran infrecuentes, esporádicos e incidentales. Vieta fue el primero en emplear letras sistemáticamente y con un propósito, no sólo para representar

---

[6] Introducción al Arte Analítico, 1591 = *Opera*, 1-12.

una incógnita o las potencias de una incógnita, sino como coeficientes generales. Habitualmente usaba consonantes para las cantidades conocidas y vocales para las desconocidas. Llamaba a su álgebra simbólica *logistica speciosa,* en oposición a *logistica numerosa.* Vieta era plenamente consciente de que cuando estudiaba la ecuación general de segundo grado $ax^2 + bx + c = 0$ (en nuestra notación), estaba estudiando toda una clase de expresiones. Al hacer la distinción entre *logistica speciosa* y *logistica numerosa* en su *Isagoge,* Vieta trazó la línea divisoria entre la aritmética y el álgebra. El álgebra, la *logistica speciosa,* dijo, es un método de operar con especies o formas de cosas. La aritmética, la *numerosa,* trata de números. Así, en un solo paso, el álgebra se convirtió en un estudio de tipos generales de formas y ecuaciones, pues lo que se hace para el caso general cubre una infinidad de casos particulares. Vieta empleaba coeficientes literales solamente para representar números positivos.

Vieta trató de establecer las identidades algebraicas ocultas en forma geométrica en las obras griegas clásicas, pero, a su juicio, claramente reconocibles en Diofanto. En efecto, como observamos en el capítulo 6, éste había realizado muchas transformaciones de expresiones algebraicas mediante identidades que no citaba explícitamente. En sus *Zeteticorum Libri Quinque,* Vieta intentó recobrar tales identidades. Completó el cuadrado de una expresión cuadrática general y expresó identidades de tipo general como

$$a^3 + 3a^2b + 3ab^2 + b^3 = (a + b)^3$$

aunque él lo escribía así:

$$\text{a cubus + b in a quadr. 3 + a in b quad. 3 + b cubo aequalia } \overline{\text{a + b}}\text{ cubo.}$$

Podría parecer lógico que los sucesores de Vieta se hubiesen impresionado inmediatamente por la idea de los coeficientes generales. Sin embargo, parece ser que la introducción de letras para denotar clases de números fue aceptada como un progreso de poca entidad en el desarrollo del simbolismo. La idea de los coeficientes literales se deslizó en las matemáticas casi por casualidad. No obstante, las ideas de Vieta sobre el simbolismo fueron apreciadas y flexibilizadas por Harriot, Girard y Oughtred.

A Descartes se deben ciertas mejoras en el uso de las letras de Vieta. Empleaba las primeras letras del alfabeto para las cantidades conocidas, y las últimas para las incógnitas, como se hace modernamente. Pero Descartes, como Vieta, sólo utilizaba las letras para representar números positivos, si bien no dudaba en efectuar restas entre términos con coeficientes literales. Hasta que lo hizo John Hudde (1633-1704), en 1657, no se empleó una letra para designar números positivos y negativos. Newton lo hacía con toda libertad.

Es preciso mencionar a Leibniz en la historia del simbolismo, aunque es posterior a los avances más significativos en el álgebra. Realizó prolongados estudios de diversas notaciones, experimentó con símbolos, pidió opinión a sus contemporáneos, y después escogió lo mejor. Encontraremos algo de su simbolismo en nuestro estudio del cálculo. Apreciaba ciertamente el gran ahorro de pensamiento que unos buenos símbolos hacen posible.

Así, hacia fines del siglo XVII, el uso deliberado del simbolismo, esto es, no incidental o accidental, y la consciencia de la potencia y generalidad que confiere, habían hecho su entrada en la matemática. Desgraciadamente, hubo demasiados símbolos introducidos a prueba y sin reflexión por parte de personas que no percibían la importancia del instrumento simbólico. Al observar esto, el historiador Florian Cajori se vio impulsado a decir que «nuestros símbolos de hoy son un mosaico de signos individuales de sistemas rechazados».

4. La solución de las ecuaciones de tercer y cuarto grados

La solución de las ecuaciones de segundo grado por el método de completar el cuadrado era conocida desde la época de los babilonios, y prácticamente el único progreso en este tema hasta 1500 fue llevado a cabo por los hindúes, que trataban ecuaciones como $x^2 + 3x + 2$ y $x^2 - 3x - 2$ como un solo tipo, mientras que sus predecesores, e incluso la mayoría de sus sucesores renacentistas, preferían tratar la última en la forma $x^2 = 3x + 2$. Cardano, como ya observamos, resolvió de hecho una ecuación de segundo grado con raíces complejas, pero despreció las soluciones como inútiles. La ecuación cúbica, excepto en casos aislados, había desafiado a la matemática hasta entonces; en fecha tan tardía como 1494, Pacioli afirmaba que la solución de las ecuaciones generales de tercer grado era imposible.

Scipione dal Ferro (1465-1526), profesor de matemáticas en Bolo-

nia, resolvió hacia 1500 ecuaciones del tipo $x^3 + mx = n$, aunque no publicó su método porque en los siglos XVI y XVII los descubrimientos solían mantenerse secretos para desafiar a los rivales a resolver el mismo problema. No obstante, hacia 1510 confió su método a Antonio María Fior (de la primera mitad del siglo XVI) y a su yerno y sucesor Annibale della Nave (1500?-58).

No sucedió nada más hasta que Niccolò Fontana de Brescia (1499?-1557) entró en escena. De niño recibió en la cara un corte de sable de un soldado francés, lo que le causó una tartamudez, a consecuencia de lo cual fue apodado Tartaglia, «tartamudo». Educado en la pobreza, se enseñó a sí mismo latín, griego y matemáticas. Se ganaba la vida dando clases de ciencias en diversas ciudades italianas. En 1535, Fior desafió a Tartaglia a resolver treinta ecuaciones de tercer grado. Tartaglia, que decía haber resuelto ya ecuaciones cúbicas de la forma $x^3 + mx^2 = n$, con m y n positivos, resolvió las treinta, incluyendo las del tipo $x^3 + mx = n$.

Presionado por Cardano a revelar su método, Tartaglia se lo dio en un oscuro verso, tras haberle prometido Cardano mantenerlo secreto. Esto sucedía en 1539. En 1542, Cardano, y su alumno Ludovico Ferrari (1522-65), con ocasión de una visita de della Nave, determinó que el método de dal Ferro era el mismo que el de Tartaglia. A pesar de su promesa, Cardano publicó su versión del método en su *Ars Magna*. En el capítulo 11 dice que «Scipio Ferro de Bolonia descubrió hace más de treinta años esta fórmula y se la dio a Antonio María Fior de Venecia, cuyo concurso con Niccolò Tartaglia de Brescia dio a éste ocasión de descubrirla. Me la dio en respuesta a mis peticiones, aunque ocultando la demostración. Armado de esta asistencia, busqué su demostración en [varias] formas. Fue muy difícil. A continuación sigue mi versión».

Tartaglia protestó por la ruptura de la promesa, y en *Quesiti ed invenzioni diverse* (1546) presentó su propia versión. Sin embargo, ni en este libro ni en su *General trattato de' numeri e misure* (1556), que es una buena presentación de los conocimientos aritméticos y geométricos de la época, ofreció nada más de la propia ecuación de tercer grado. La disputa sobre quién resolvió primero la ecuación de tercer grado condujo a un conflicto abierto entre Tartaglia y Ferrari, que se extendió entre salvajes peleas y en el cual Cardano no tomó parte. El propio Tartaglia no estaba por encima de los reproches; publicó una traducción de parte de la obra de Arquímedes, que en realidad había copiado de Guillermo de Moerbecke (m. *c.* 1281), y pretendía haber

descubierto la ley del movimiento de un objeto en un plano inclinado... que en realidad era de Jordanus Nemorarius.

Cardano ilustra su método con la ecuación $x^3 + 6x = 20$, pero nosotros, para comprobar la generalidad del mismo, consideraremos

$$x^3 + mx = n \tag{4}$$

con m y n positivos. Cardano introduce dos cantidades t y u, imponiendo que se verifiquen las dos relaciones siguientes:

$$t - u = n \tag{5}$$

y

$$(tu) = \left(\frac{m}{3}\right)^3. \tag{6}$$

Afirma entonces que

$$x = \sqrt[3]{t} - \sqrt[3]{u}. \tag{7}$$

Por eliminación en (5) y (6), y resolviendo la ecuación de segundo grado resultante, obtiene

$$t = \sqrt{\left(\frac{n}{2}\right)^2 + \left(\frac{m}{3}\right)^3} + \frac{n}{2}, \qquad u = \sqrt{\left(\frac{n}{2}\right)^2 + \left(\frac{m}{3}\right)^3} - \frac{n}{2}. \tag{8}$$

Como Cardano, hemos tomado la raíz positiva. Una vez calculados t y u, Cardano toma la raíz cúbica de cada uno y por (7) obtiene un valor de x. Esta es probablemente la misma raíz obtenida por Tartaglia.

Este es el método de Cardano. Tenía que probar, sin embargo, que (7) da un valor correcto de x. Su prueba es geométrica; para Cardano, t y u son volúmenes de cubos de lados $\sqrt[3]{t}$ y $\sqrt[3]{u}$ respectivamente, y el producto $\sqrt[3]{t} \cdot \sqrt[3]{u}$ representa un rectángulo formado por ambos lados, cuya área es $m/3$. Igualmente, cuando nosotros expresamos $t - u = n$, Cardano dice que la diferencia de ambos volúmenes es n. A continuación afirma que la solución x es la diferencia de los lados de los dos cubos, es decir, $x = \sqrt[3]{t} - \sqrt[3]{u}$. Para demostrar

que este valor de x es correcto, enuncia y prueba el lema geométrico siguiente:

FIGURA 13.2

Si de un segmento AC (fig. 13.2) se elimina un segmento BC, entonces el cubo de lado AB es igual al cubo de lado AC menos el cubo de lado BC menos el triple del paralelepípedo rectángulo de aristas AC, AB y BC. Este lema geométrico, evidentemente, no dice más que

$$(\sqrt[3]{t} - \sqrt[3]{u})^3 = t - u - 3(\sqrt[3]{t} - \sqrt[3]{u})\sqrt[3]{t}\,\sqrt[3]{u}. \qquad (9)$$

Dando por demostrado este lema (que puede comprobarse por la expresión del cubo de un binomio, pero que Cardano establece citando teoremas de Euclides), sólo tiene que observar que si $x = \sqrt[3]{t} - \sqrt[3]{u}$, $t - u = n$ y $\sqrt[3]{t}\,\sqrt[3]{u} = \frac{m}{3}$, lo que afirma el lema es $x^3 = n - mx$. Por ello, si elige t y u de forma que satisfagan las condiciones (5) y (6), el valor de x dado por (7) en términos de t y u satisface la ecuación de tercer grado. Da entonces una regla aritmética puramente verbal para explicar el método, que nos dice que debemos formar el término $\sqrt[3]{t} - \sqrt[3]{u}$, donde t y u vienen dados en (8) en términos de m y n.

Cardano, como Tartaglia, también resuelve ecuaciones particulares de los tipos

$$x^3 = mx + n, \quad x^3 + mx + n = 0, \quad x^3 + n = mx.$$

Ha de tratar cada uno de dichos casos separadamente, y todos ellos independientemente de la ecuación (4), ya que, en primer lugar, hasta esta época las ecuaciones que escribían los europeos sólo contenían términos con números positivos, y, en segundo lugar, Cardano tenía que dar una justificación geométrica independiente para cada caso.

Cardano también muestra cómo resolver ecuaciones como $x^3 + 6x^2 = 100$. Sabía cómo eliminar el término en x^2; como el coeficiente es 6, sustituye x por $y - 2$ y obtiene $y^3 = 12y + 84$. También observó que una ecuación como $x^6 + 6x^4 = 100$ puede tratarse como una

cúbica haciendo $x^2 = y$. A lo largo del libro da raíces positivas y negativas, a pesar de llamar «ficticios» a los números negativos. No tenía en cuenta, sin embargo, las raíces complejas. De hecho, en el capítulo 37, llama «falsos» a los problemas que dan lugar a raíces que no son ni verdaderas ni falsas (ni positivas ni negativas). El libro es detallista, incluso aburrido para el lector moderno, y ello porque Cardano trata separadamente toda una multitud de casos, no sólo de la ecuación de tercer grado, sino también de las ecuaciones cuadráticas auxiliares que ha de resolver para hallar t y u. En cada caso, escribe la ecuación en forma tal que los coeficientes de los términos sean positivos.

Hay una dificultad con la solución de Cardano de la ecuación de tercer grado que él observó, pero no resolvió. Cuando todas las raíces de la ecuación son reales y distintas, se puede probar que t y u son complejos, ya que el radicando en (8) es negativo; y sin embargo, necesitamos $\sqrt[3]{t}$ y $\sqrt[3]{u}$ para obtener x. Esto significa que ciertos números reales pueden expresarse en términos de raíces cúbicas de números complejos. Sin embargo, esas tres raíces reales no pueden obtenerse por medios algebraicos, es decir, por radicales. Tartaglia llamó a este caso «irreducible». Uno pensaría que la posibilidad de expresar números reales como combinaciones de números complejos hubiera persuadido a Cardano a considerarlos seriamente, pero no fue así.

Vieta, en *De Aequationum Recognitione et Emendatione*, escrito en 1591 y publicado en 1615 [8], pudo resolver el caso irreducible de la ecuación de tercer grado empleando una identidad trigonométrica, evitando así la fórmula de Cardano. Este método se utiliza hoy día. Empieza por la identidad

$$\cos 3A = 4\cos^3 A - 3\cos A. \qquad (10)$$

Haciendo $z = \cos A$, esta identidad se transforma en

$$z^3 - \frac{3}{8}z - \frac{1}{4}\cos 3A = 0. \qquad (11)$$

Supongamos que la ecuación dada es (Vieta trabajó con $x^3 - 3a^2x = a^2b$, con $a > b/2$)

[8] De la Revisión y Corrección de Ecuaciones, *Opera*, 82-162.

$$y^3 + pq + q = 0. \quad (12)$$

Introduciendo $y = nz$, donde n está a nuestra disposición, podemos hacer coincidir los coeficientes de (12) con los de (11). Sustituyendo $y = nz$ en (12) resulta

$$z^3 + \frac{p}{n^2} z + \frac{q}{n^3} = 0. \quad (13)$$

Ahora imponemos a n que $p/n^2 = -3/4$, de forma que

$$n = \sqrt{-4p/3}. \quad (14)$$

Con esta elección de n, tomamos un valor A tal que

$$\frac{q}{n^3} = -\frac{1}{4} \cos 3A \quad (15)$$

o bien

$$\cos 3A = -\frac{4q}{n^3} = \frac{-q/2}{\sqrt{-p^3/27}}. \quad (16)$$

Se puede demostrar que si las tres raíces son reales, entonces p es negativo de tal modo que n es real. Además, puede verse que $|\cos A| < 1$. Por tanto, puede hallarse $3A$ con una tabla.

Cualquiera que sea el valor de A, $\cos A$ satisface (11) por ser una identidad. Ahora bien, se ha elegido A de forma que (13) sea un caso particular de (11). Para este valor de A, $\cos A$ satisface (13). Pero el valor de A está determinado por (16), lo cual fija $3A$, y para cualquier A que satisfaga (16), también la satisfacen $A + 120°$ y $A + 240°$, y como $z = \cos A$, hay, por tanto, tres valores que satisfacen (13):

$$\cos A, \cos (A + 120°) \text{ y } \cos (A + 240°).$$

Los tres valores que satisfacen (12) son el producto de n por estos valores de z, donde n está dado en (14). Vieta obtuvo sólo una raíz.

La ecuación de tercer grado tiene, desde luego, tres raíces. Fue Leonhard Euler, en 1732, el primero en ofrecer una discusión comple-

ta de la solución de Cardano de la ecuación cúbica, insistiendo en que siempre hay tres raíces y especificando cómo se calculan [9]. Si ω y ω^3 son las raíces complejas de $x^3 - 1 = 0$, es decir, las raíces de $x^2 + x + 1 = 0$, entonces las tres raíces cúbicas de t y u en (8) son

$$\sqrt[3]{t},\ \omega\sqrt[3]{t},\ \omega^2\sqrt[3]{t} \text{ y } \sqrt[3]{u},\ \omega\sqrt[3]{u},\ \omega^2\sqrt[3]{u}.$$

Debemos ahora escoger un miembro del primer conjunto y otro del segundo de forma que el producto sea el número real $m/3$ (ver la ecuación (6) en la solución de Cardano). Como ω y ω^3 son raíces de la unidad, $\omega \cdot \omega^3 = \omega^3 = 1$ y las opciones válidas para x, a la vista de (7), son

$$x_1 = \sqrt[3]{t} - \sqrt[3]{u},\ x_2 = \omega\sqrt[3]{t} - \omega^2\sqrt[3]{u},\ x_3 = \omega^2\sqrt[3]{t} - \omega\sqrt[3]{u}. \qquad (17)$$

Al éxito en la resolución de la ecuación de tercer grado sucedió casi inmediatamente la solución de la de cuarto. El método se debe a Ludovico Ferrari, y fue publicado en el *Ars Magna* de Cardano. Aquí lo describiremos en notación moderna y con coeficientes literales para mostrar su generalidad. La ecuación es

$$x^4 + bx^3 + cx^2 + dx + e = 0. \qquad (18)$$

Transponiendo obtenemos

$$x^4 + bx^3 = -cx^2 - dx - e. \qquad (19)$$

Completamos ahora el cuadrado en el primer miembro sumando $\left(\frac{1}{2}bx\right)^2$:

$$\left(x^2 + \frac{1}{2}bx\right)^2 = \left(\frac{1}{4}b^2 - c\right)x^2 - dx - e. \qquad (20)$$

Sumemos ahora $\left(x^2 + \frac{1}{2}bx\right)y + \frac{1}{4}y^2$ a ambos miembros, y obtendremos

[9] *Comm. Acad. Sci. Petrop.*, 6, 1732/33, 217-31, pub. 1738 = *Opera* (1), 6, 1-19.

$$\left(x^2 + \frac{1}{2}bx\right)^2 + \left(x^2 + \frac{1}{2}bx\right)y + \frac{1}{4}y^2 \tag{21}$$
$$= \left(\frac{1}{4}b^2 - c + y\right)x^2 + \left(\frac{1}{2}by - d\right)x + \frac{1}{4}y^2 - e.$$

Igualando a cero el discriminante de la expresión de segundo grado en x del segundo miembro, podemos convertir éste en el cuadrado perfecto de una expresión de primer grado en x. Hacemos, pues,

$$\left(\frac{1}{2}by - d\right)^2 - 4\left(\frac{1}{4}b^2 - c + y\right)\left(\frac{1}{4}y^2 - e\right) = 0. \tag{22}$$

Esta es una ecuación de *tercer grado* en y. Elijamos una raíz cualquiera de esta ecuación cúbica y sustituyámosla por y en (21). Usando el hecho de que el primer miembro es también un cuadrado perfecto, obtenemos una ecuación de segundo grado en x igualando cualquier par de funciones lineales de x, una opuesta de la otra. La resolución de estas dos ecuaciones cuadráticas nos da 4 raíces para x. Si se elige otra raíz de (22) se obtiene otra ecuación distinta en (21), pero las mismas cuatro raíces.

Para presentar el método de Ferrari, Cardano, en el capítulo 39 de la *Ars Magna*, resuelve una multitud de casos especiales, todos con coeficientes numéricos. Así, resuelve ecuaciones de los tipos

$$x^4 = bx^2 + ax + n, \qquad x^4 = bx^2 + cx^3 + n$$
$$x^4 = cx^3 + n \qquad x^4 = ax + n$$

Da, como en el caso de la ecuación de tercer grado, una prueba geométrica de los pasos algebraicos básicos, y después da la regla de solución en palabras.

A base de resolver numerosos ejemplos de ecuaciones de tercer y cuarto grado, Cardano, Tartaglia y Ferrari dieron prueba de haber buscado y obtenido métodos que funcionaban para todos los casos de los grados respectivos. El interés en la generalidad es una característica nueva. Su trabajo precedió a la introducción de los coeficientes literales por parte de Vieta, de forma que no pudieron beneficiarse de tal instrumento. Vieta, que ya había hecho posible la generalidad en la

demostración mediante la introducción de los coeficientes literales, buscaba ahora otro tipo de generalidad. Observó que los métodos de resolución de las ecuaciones de segundo, tercer y cuarto grados eran muy diferentes. Buscó, por tanto, un método que fuese válido para las ecuaciones de cualquier grado. Su primera idea fue eliminar el término de grado inmediatamente inferior al máximo mediante una sustitución. Tartaglia había hecho esto para la ecuación cúbica, pero no lo intentó para todas las ecuaciones.

En el *Isagoge,* Vieta hace lo siguiente. Para resolver la ecuación de segundo grado

$$x^2 + 2bx = c$$

hace

$$x + b = y.$$

Entonces

$$y^2 = x^2 + 2bx + b^2.$$

A la vista de la ecuación original,

$$y = \sqrt{c + b^2}.$$

Por tanto,

$$x = y - b = \sqrt{c + b^2} - b.$$

En el caso de la ecuación de tercer grado

$$x^3 + bx^2 + cx + d = 0.$$

Vieta empieza haciendo $x = y - b/3$. Esta sustitución da lugar a la cúbica reducida

$$y^3 + py + q = 0. \qquad (23)$$

A continuación introduce una transformación más, que, de hecho, es la que aún se estudia hoy día. Define

$$y = z - \frac{p}{3z} \tag{24}$$

y obtiene

$$z^3 - \frac{p^3}{27z^3} + q = 0.$$

Resuelve entonces la ecuación cuadrática en z^3 obteniendo

$$z^3 = -\frac{q}{2} \pm \sqrt{R}, \quad \text{donde} \quad R = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2.$$

Aquí, como en el método de Cardano, z^3 puede tener dos valores. Aunque Vieta sólo empleaba la raíz cúbica positiva de z^3, se puede considerar las seis raíces (complejas), y (24) mostraría que resultan sólo tres valores distintos de y a partir de los seis valores de z.

Para resolver la ecuación general de cuarto grado

$$x^4 + bx^3 + cx^2 + dx + e = 0,$$

Vieta hace $x = y - b/4$, reduciendo la ecuación a

$$x^4 + px^2 + qx + r = 0.$$

Transpone ahora los tres últimos términos y suma $2x^2y^2 + y^4$ a ambos miembros. Esto convierte al primer miembro en cuadrado perfecto y, como en el método de Ferrari, eligiendo y adecuadamente consigue expresar el segundo miembro como un cuadrado perfecto de la forma $(Ax + B)^2$. La elección apropiada de y se efectúa aplicando la condición del discriminante a una ecuación de segundo grado, lo cual conduce a una ecuación de sexto grado en y que, afortunadamente, es de tercer grado en y^2. Este paso y el resto del trabajo son idénticos a los del método de Ferrari.

Otro método general explorado por Vieta es la descomposición del polinomio en factores de primer grado, como, por ejemplo, $x^2 + 5x + 6 = (x + 2)(x + 3)$. No tuvo éxito en ello, en parte por rechazar las raíces que no fuesen positivas, y en parte por no poseer una teoría suficiente, como el teorema del factor, en la cual basar un método

general. Thomas Harriot tuvo la misma idea y fracasó por las mismas razones.

La búsqueda de métodos algebraicos generales se desplazó a continuación hacia la resolución de ecuaciones de grado mayor que cuatro. James Gregory, que había proporcionado métodos propios de resolución de la ecuación de tercer y cuarto grados, trató de emplearlos en la solución de la de quinto. El y Ehrenfried Walter von Tschirnhausen (1651-1708) probaron transformaciones para reducir las ecuaciones de orden superior a dos términos, una potencia de x y una constante. Estos métodos de solución de ecuaciones de grado mayor que cuatro fracasaron. En posteriores trabajos sobre integración, Gregory dio por sentado que no era posible resolver algebraicamente la ecuación general de grado n para $n > 4$.

5. La teoría de ecuaciones

El trabajo sobre los métodos de resolución de ecuaciones produjo una serie de teoremas y observaciones que se estudian todavía hoy en la teoría elemental de ecuaciones. La cuestión del número de raíces que puede tener una ecuación fue objeto de atención. Cardano había introducido las raíces complejas, y por un tiempo pensó que una ecuación podía tener cualquier número de raíces, pero pronto se dio cuenta de que una ecuación de tercer grado tiene 3 raíces, una de cuarto 4, y así sucesivamente. En *L'Invention nouvelle*, Albert Girard infiere y enuncia que una ecuación polinómica de grado n tiene n raíces si se cuentan las raíces imposibles (es decir, complejas) y si se tienen en cuenta las repetidas. Girard, sin embargo, no dio prueba alguna. Descartes, en el tercer libro de *La Géometrie*, dice que una ecuación puede tener tantas raíces como el número de dimensiones (el grado) de la incógnita, usando la expresión «puede tener» por considerar las raíces negativas como falsas. Más tarde, al incluir las raíces imaginarias y las negativas a efectos de contar las raíces, concluyó que hay tantas como indica el grado.

La siguiente cuestión importante fue la de averiguar el número de raíces positivas, negativas y complejas. Cardano observó que las raíces complejas de una ecuación (con coeficientes reales) se dan por pares. Newton lo demostró en su *Arithmetica Universalis*. Descartes, en *La Géometrie*, enunció sin demostración la regla de los signos, conocida como «regla de Descartes», que afirma que el máximo número de

raíces positivas de $f(x) = 0$, donde f es un polinomio, es el número de variaciones del signo de los coeficientes, y que el máximo número de raíces negativas es el número de apariciones de dos signos «+» o dos signos «−» consecutivamente. En terminología actual, la última parte de la regla afirma que el número máximo de raíces negativas es el número de variaciones en la ecuación $f(-x) = 0$. Esta regla fue demostrada por varios matemáticos del siglo XVIII. La prueba que se da normalmente en la actualidad se debe a Abbé Jean-Paul de Gua de Malves (1712-85), que también demostró que la ausencia de $2m$ términos consecutivos indica que hay $2m + 2$ ó $2m$ raíces complejas, según que los dos términos entre los que se halla la deficiencia tengan signos iguales o distintos.

En su *Arithmetica Universalis,* Newton describió, sin prueba, otro método para determinar el número máximo de raíces positivas y negativas y, con ello, el mínimo número de raíces complejas. Su método es más complicado de aplicar, pero da mejores resultados que la regla de los signos de Descartes. Fue finalmente demostrado como caso particular de un teorema más general de Sylvester [10]. Un poco antes, Gauss había demostrado que si el número de raíces positivas queda por debajo del número de variaciones del signo, tal diferencia debe ser un número par.

Otra clase de resultados se refiere a las relaciones entre las raíces y los coeficientes de una ecuación. Cardano descubrió que la suma de las raíces es el opuesto del coeficiente de x^{n-1}, que la suma de los productos de dos en dos es el coeficiente de x^{n-2}, etc. Tanto Cardano como Vieta (en *De Aequationum Recognitione et Emendatione*) emplearon la primera de estas relaciones entre raíces y coeficientes de ecuaciones de grado pequeño para eliminar el término en x^{n-1} en las ecuaciones polinómicas de la forma que ya hemos descrito. Newton enunció la relación entre raíces y coeficientes en su *Arithmetica Universalis,* y también James Gregory en una carta a John Collins (1625-83), secretario de la Royal Society, aunque ninguno de ellos dio ninguna demostración.

Vieta y Descartes construyeron ecuaciones cuyas raíces fuesen mayores o menores que las de otra ecuación dada previamente. El proceso consiste simplemente en reemplazar x por $y + m$. Ambos utilizaron la transformación $y = mx$ para obtener una ecuación cuyas raíces fuesen el producto de m por las de la ecuación dada. Para

---

[10] *Proc. London Math. Soc.*, 1, 1865, 1-16 = *Math. Papers*, 2, 498-513.

Descartes, el primero de estos procesos tenía la significación antes apuntada de que las raíces falsas (negativas) pudiesen hacerse verdaderas (positivas) y recíprocamente.

Descartes demostró también que si una ecuación de tercer grado con coeficientes racionales tiene una raíz racional, entonces el polinomio puede expresarse como producto de factores con coeficientes racionales.

Otro resultado importante es el actualmente conocido como «teorema del factor». En el tercer libro de *La Géometrie*, Descartes enuncia que $f(x)$ es divisible por $x - a$, con a positivo, si y sólo si a es una raíz de $f(x) = 0$, y por $x + a$ si y sólo si a es una raíz falsa. Con éste y otros resultados, Descartes establece el método moderno de hallar las raíces racionales de una ecuación polinómica. Tras convertir en 1 el coeficiente de mayor grado, hace que todos los coeficientes sean enteros multiplicando las raíces de la ecuación por el factor preciso, mediante la regla por él dada de sustituir x por y/m en la ecuación. Las raíces racionales de la ecuación original deben ser ahora factores enteros del término constante de la nueva ecuación. Si, a base de probar, se encuentra una raíz a, por el teorema del factor, $y - a$ es divisor del nuevo polinomio en y. Descartes observa entonces que eliminando este factor se reduce el grado de la ecuación y se puede trabajar con la ecuación reducida.

Newton, en *Arithmetica Universalis*, y otros anteriormente, enunciaron resultados de acotación superior de las raíces de una ecuación. Uno de estos teoremas tiene que ver con el cálculo, y será enunciado en el capítulo 17 (sec. 7). Newton descubrió la relación entre las raíces y el discriminante de la ecuación, por ejemplo, que $ax^2 + bx + c = 0$ tiene raíces iguales, reales o no reales según que $b^2 - 4ac$ sea igual, mayor o menor que 0.

En *La Géometrie*, Descartes introduce el principio de coeficientes indeterminados, que puede ilustrarse del siguiente modo: para descomponer $x^2 - 1$ en dos factores linales, se expresa

$$x^2 - 1 = (x + b)(x + d).$$

Haciendo la multiplicación en el segundo miembro e igualando los coeficientes de las potencias de x en ambos miembros, se obtiene

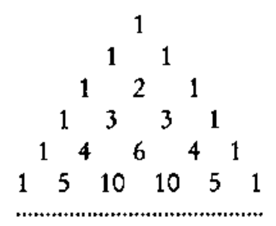
$$b + d = 0$$
$$bd = -1,$$

de donde se obtienen *b* y *d*. Descartes insistió en la utilidad de este método.

Otro método, el de inducción matemática, hizo su entrada explícitamente en el álgebra a finales del siglo XVI. Desde luego, el método se halla implícito incluso en la demostración de Euclides de la infinitud de los números primos. En ella prueba que si hay *n* números primos, debe haber $n + 1$, y como hay un primer número primo, el número de ellos debe ser infinito. El método fue explícitamente reconocido por Maurolico en su *Arithmetica* de 1575, y lo empleó para probar, por ejemplo, que $1 + 3 + 5 + \ldots + (2n - 1) = n^2$. Pascal comentó en una carta su conocimiento de la introducción del método por parte de Maurolico y lo usó en su *Traité du triangle arithmétique* (1665), donde presenta lo que hoy llamamos triángulo de Pascal (sec. 6).

6. El teorema binomial y cuestiones afines

El teorema binomial (o fórmula del binomio) para exponentes enteros positivos, es decir, el desarrollo de $(a + b)^n$ para *n* entero positivo, ya era conocido por los árabes del siglo XIII. Hacia 1544, Stifel introdujo el término «coeficiente binomial» y mostró cómo calcular $(1 + a)^n$ a partir de $(1 + a)^{n-1}$. El cuadro de números



en el cual cada número es la suma de los dos que están inmediatamente sobre él, ya conocido de Tartaglia, Stifel y Stevin, fue empleado por Pascal (1654) para obtener los coeficientes del desarrollo del binomio. Por ejemplo, los números de la cuarta fila son los coeficientes del desarrollo de $(a + b)^3$. A pesar de que este cuadro numérico era conocido por muchos de sus predecesores, ha terminado por llamarse «triángulo de Pascal».

En 1665, Newton mostró cómo calcular $(1 + a)^n$ directamente, sin hacer referencia a $(1 + a)^{n-1}$. Se convenció entonces de que el

desarrollo era válido para valores fraccionarios y negativos de n (resultando una serie infinita en este caso), y enunció, aunque nunca demostró, tal generalización. Sí llegó a verificar que la serie de $(1 + x)^{1/2}$ multiplicada por sí misma da $1 + x$, pero ni él ni James Gregory (que llegó al mismo teorema independientemente) creyeron necesaria una demostración. En dos cartas, del 6 de junio y del 4 de octubre de 1676, a Henry Oldenburg (*c.* 1615-77), secretario de la Royal Society, Newton enunció el resultado más general que conocía en 1669, a saber, el desarrollo de $(P + PQ)^{m/n}$. Lo consideraba un método útil para extraer raíces, pues si Q es menor que 1 (una vez sacado P factor común), los términos sucesivos, al ser potencias de Q, tiene valores cada vez menores.

Con independencia del trabajo sobre el teorema del binomio, las fórmulas para el número de permutaciones y de combinaciones de n objetos tomados r a r habían aparecido en las obras de una serie de matemáticos, entre ellos Bhāskara y el francés Levi ben Gerson (1321). Pascal observó que la fórmula de las combinaciones, denotada generalmente por ${}_nC_r$ o $\binom{n}{r}$, proporciona también los coeficientes binomiales. Es decir, para n fijo y r entre 0 y n, la fórmula da los sucesivos coeficientes. Bernoulli, en su *Ars Conjectandi* (1713), extendió la teoría de combinaciones y demostró el teorema binomial en el caso de n entero positivo mediante las fórmulas de las combinaciones.

Las investigaciones sobre permutaciones y combinaciones están conectadas con otro desarrollo, la teoría de la probabilidad, que iba a alcanzar importancia fundamental en el siglo XIX, pero apenas merece mención en los siglos XVI y XVII. El problema de la probabilidad de sacar un número concreto al lanzar dos dados se había planteado ya en la época medieval. Otro problema, el de cómo dividir las ganancias entre dos jugadores cuando éstas corresponden al primer jugador que obtenga n puntos, y el juego resulta interrumpido cuando el primero lleva ganados p puntos y el segundo q, aparece en la *Summa* de Pacioli y en libros de Cardano, Tartaglia y otros. Este problema adquirió cierta importancia cuando, propuesto a Pascal por Antoine Gombaud, Chevalier de Méré (1610-85), Pascal y Fermat se cartearon al respecto. El problema y su solución carecen de importancia, pero el trabajo de ambos marca el comienzo de la teoría de la probabilidad. Los dos aplicaron para ello la teoría de combinaciones.

El primer libro de significación sobre la probabilidad es el *Ars*

Conjectandi de Bernoulli. Su aportación más importante, aún llamada «teorema de Bernoulli», dice que si p es la probabilidad de un suceso y q es la probabilidad de que éste no se dé, entonces la probabilidad de que dicho suceso se dé al menos m veces en n intentos es la suma de los términos del desarrollo de $(p + q)^n$ desde p^n hasta el término que contiene $p^m q^{m-n}$.

7. La teoría de números

Mientras que los intereses prácticos fueron el estímulo para el progreso en el cálculo, el simbolismo y la teoría de ecuaciones, fue la preocupación por problemas puramente matemáticos la que condujo a una actividad renovada en la teoría de números. El tema, desde luego, había sido iniciado por los griegos clásicos, y Diofanto había añadido la parte relativa a las ecuaciones indeterminadas. Hindúes y árabes evitaron que cayese en el olvido, y aunque casi todos los algebristas del Renacimiento que hemos mencionado hicieron conjeturas y observaciones, el primer europeo que hizo extensas e impresionantes contribuciones a la teoría de números y dio al tema un enorme impulso fue Pierre de Fermat (1601-65).

Nacido en una familia de comerciantes, se formó como abogado en la ciudad francesa de Toulouse, ganándose la vida con dicha profesión. Durante cierto período fue consejero del parlamento de Toulouse. Aunque las matemáticas eran una afición para Fermat, y sólo podía dedicarles escaso tiempo, contribuyó con resultados de primera importancia a la teoría de números y al cálculo, fue uno de los dos creadores de la geometría de coordenadas y, como ya hemos visto, inició junto a Pascal la investigación sobre la probabilidad. Como todos los matemáticos de su siglo, trabajó en problemas de las ciencias y dejó una duradera contribución a la óptica: el «principio del tiempo mínimo de Fermat» (cap. 24, sec. 3). La mayoría de los resultados de Fermat nos son conocidos a través de cartas escritas a sus amigos. Publicó sólo unos pocos artículos, y algunos de sus libros y artículos fueron publicados tras su muerte.

Fermat consideraba que se había descuidado la teoría de números. Se quejó en una ocasión de que apenas nadie proponía o entendía cuestiones aritméticas y se preguntó: «¿se debe a que hasta ahora la aritmética ha sido tratada más geométrica que aritméticamente?» El

propio Diofanto, observó, estaba atado en cierta medida a la geometría. Para Fermat, la aritmética tiene un dominio que le es propio, la teoría de los números enteros.

La obra de Fermat en teoría de números determinó la dirección del trabajo en el área hasta las contribuciones de Gauss. El punto de partida de Fermat fue Diofanto. Muchas traducciones de su *Arithmetica* habían sido realizadas por matemáticos renacentistas. En 1621, Bachet de Méziriac publicó el texto griego y una traducción latina. Esta fue la edición que manejó Fermat, anotando la mayoría de sus resultados en los márgenes, y comunicando unos cuantos a sus amigos por carta. El ejemplar que contiene las notas marginales de Fermat fue publicado por su hijo en 1670.

Fermat enunció muchos teoremas sobre teoría de números, pero sólo en un caso dio una demostración, y ésta apenas esbozada. Los mejores matemáticos del siglo XVIII dedicaron grandes esfuerzos a probar sus resultados (cap. 25, sec. 4). Todos ellos resultaron ser correctos, excepto un error (del que luego hablaremos) y un famoso y aún no demostrado «teorema», acerca de cuya certeza todas las indicaciones son favorables. No hay duda de que poseía una gran intuición, pero no es probable que tuviese demostraciones de todas sus afirmaciones.

Un documento descubierto en 1879 entre los manuscritos de Huygens incluye un famoso método, llamado del «descenso infinito», introducido y empleado por Fermat. Para comprenderlo, consideremos el teorema enunciado por Fermat en una carta a Marin Mersenne (1588-1648) del 25 de diciembre de 1640, que afirma que un número primo de la forma $4n + 1$ puede expresarse de una y sólo una manera como suma de dos cuadrados. Por ejemplo, $17 = 16 + 1$ y $29 = 25 + 4$. El método consiste en probar que si existe un número primo de la forma $4n + 1$ que *no* posee la propiedad en cuestión, tiene que haber un número primo menor de la forma $4n + 1$ que tampoco la posea. Así, como n es arbitrario, debe haber otro más pequeño aún. Descendiendo a lo largo de todos los valores enteros positivos de n, debemos llegar a $n = 1$, o sea, al número primo $4 \cdot 1 + 1$, es decir, 5, que, por tanto, no puede gozar de la mencionada propiedad. Pero 5 sí es expresable como suma de dos cuadrados de una manera única, de donde concluimos que todo primo de la forma $4n + 1$ también lo es. Fermat envió este esbozo de demostración a su amigo Pierre de Carcavi (m. 1684) en 1659. Afirmó que había utilizado el método para probar el teorema recién descrito, pero jamás se ha hallado su

demostración. También dijo haber probado otros teoremas por este mismo método.

El método del descenso infinito es diferente de la inducción matemática. En primer lugar, no exige mostrar un caso en el que el teorema en cuestión se satisfaga, ya que el argumento puede completarse observando que el caso $n = 1$ lleva simplemente a una contradicción con algún otro hecho conocido. Además, aceptada la hipótesis para un valor de n, el método muestra que existe otro menor, pero no necesariamente el anterior, para el cual la hipótesis es cierta. Finalmente, el método demuestra la falsedad de ciertas afirmaciones, siendo, de hecho, más útil para este propósito.

Fermat afirmó también que ningún número primo de la forma $4n + 3$ puede expresarse como suma de dos cuadrados. En una nota escrita en su ejemplar de Diofanto y en la carta a Mersenne, Fermat generaliza la conocida relación triangular entre 3, 4 y 5 enunciando los siguientes teoremas: un número primo de la forma $4n + 1$ es la hipotenusa de uno y sólo un triángulo rectángulo de lados enteros. El cuadrado de $(4n + 1)$ es la hipotenusa de dos y sólo dos de tales triángulos rectángulos; su cubo, de tres; su bicuadrado, de cuatro, y así sucesivamente, hasta el infinito. Consideremos, por ejemplo, el caso $n = 1$. Entonces $4n + 1 = 5$ y 3, 4 y 5 son los lados del único triángulo rectángulo que tiene 5 como hipotenusa. Sin embargo, 5^2 es la hipotenusa de dos y sólo dos triángulos rectángulos, a saber, aquellos cuyos lados miden 15, 20, 25 y 7, 24, 25. Igualmente, 5^3 es la hipotenusa únicamente de los tres triángulos rectángulos de lados 75, 100, 125; 35, 120, 125 y 44, 117, 125.

En la carta a Mersenne, Fermat declara que el mismo número primo $4n + 1$ y su cuadrado se pueden expresar cada uno de manera única como la suma de dos cuadrados; su cubo y bicuadrado, de dos formas cada uno; sus potencias quinta y sexta, de tres, y así sucesivamente. Así, para $n = 1$, $5 = 4 + 1$ y $5^2 = 9 + 16$; $5^3 = 4 + 121 = 25 + 100$, etc. Continúa la carta: si un número primo que es la suma de dos cuadrados se multiplica por otro número primo que es también suma de dos cuadrados, el producto será la suma de dos cuadrados de dos formas distintas. Si el primer número primo se multiplica por el cuadrado del segundo, el producto será la suma de dos cuadrados de tres formas; si se multiplica por el cubo del segundo, el producto será la suma de dos cuadrados de cuatro formas, y así sucesivamente hasta el infinito.

Fermat enunció muchos teoremas sobre la representación de los

números primos en las formas $x^2 + 2y^2$, $x^2 + 3y^2$, $x^2 + 5y^2$, $x^2 - 2y^2$, y otras formas, que son extensiones de la representación como suma de dos cuadrados. Por ejemplo, todo primo de la forma $6n + 1$ puede representarse como $x^2 + 3y^2$; todo primo de la forma $8n + 1$ u $8n + 3$, como $x^2 + 2y^2$. Un número primo impar (todos menos el 2) puede expresarse de manera única como diferencia de dos cuadrados.

Dos de los teoremas enunciados por Fermat han dado en ser conocidos como el «pequeño» y el «gran», siendo éste también llamado el «último teorema de Fermat». El teorema pequeño, comunicado en carta de 18 de octubre de 1640 a su amigo Bernard Frénicle de Bessy (1605-75), afirma que si p es un número primo y a es primo con p, entonces $a^p - a$ es divisible por p.

El gran «teorema» de Fermat, que él creía haber probado, afirma que para $n > 2$ no hay soluciones enteras de $x^n + y^n = z^n$. Este teorema fue enunciado por él en una nota al margen de su libro de Diofanto, junto al problema (de Diofanto) de dividir un cuadrado dado en (suma de) dos cuadrados. Fermat añade: «Por otra parte, es imposible separar un cubo en dos cubos, un bicuadrado en dos bicuadrados, o, en general, una potencia que no sea un cuadrado en dos potencias con el mismo exponente. He descubierto una demostración verdaderamente maravillosa de esto, pero el margen no es suficiente para contenerla.» Desgraciadamente, la prueba de Fermat, si es que la obtuvo, nunca fue hallada, y cientos de los mejores matemáticos han sido incapaces de demostrar el teorema. Fermat afirmó en carta a Carvavi que había empleado el método del descenso infinito para demostrar el caso $n = 4$, pero no dio detalles. Frénicle, con las pocas indicaciones de Fermat, dio una demostración para el mencionado caso en su obra de publicación póstuma *Traité des triangles rectangles en nombres* (Tratado sobre propiedades numéricas de los triángulos rectángulos) [11].

Anticipándonos un poco, digamos que Euler demostró el teorema para $n = 3$ (cap. 25, sec. 4). Como el teorema es cierto para $n = 3$, es cierto para cualquier múltiplo de 3, pues si no lo fuese para $n = 6$, por ejemplo, habría enteros x, y, z tales que

$$x^6 + y^6 = z^6$$

[11] *Mém. de l'Acad. des Sci., Paris,* 5, 1729, 83-166.

Pero entonces

$$(x^2)^3 + (y^2)^3 = (z^2)^3$$

y el teorema sería falso para $n = 3$. Por tanto, sabemos que el teorema de Fermat es cierto para un número infinito de valores de n, pero no sabemos todavía si lo es para todos. En realidad sólo se necesita probarlo para $n = 4$ y para n primo impar. En efecto, supongamos primero que n no es divisible por un primo impar; debe ser entonces una potencia de 2, y al ser mayor que 2 tiene que ser 4 o divisible por 4. Sea $n = 4m$. Entonces la ecuación $x^n + y^n = z^n$ se convierte en

$$(x^m)^4 + (y^m)^4 = (z^m)^4$$

Si el teorema no fuese cierto para n, tampoco lo sería para $n = 4$. Luego si es cierto para $n = 4$, lo es para todo n no divisible por un número primo impar. Si $n = pm$, donde p es primo impar, entonces, si el teorema no fuese cierto para n, tampoco lo sería para el exponente p. Luego si es cierto para $n = p$, lo es para todo n divisible por un número primo impar.

Fermat cometió algunos errores. Creía haber hallado una solución al viejo problema de construir una fórmula que diese números primos para todos los valores de la variable n. Sin embargo, no es difícil demostrar que $2^m + 1$ no puede ser primo a no ser que m sea una potencia de 2. En muchas cartas, datadas de 1640 en adelante [12], Fermat afirmó el recíproco (es decir, que $(2)^{2^n} + 1$ representa una serie de números primos), aunque admitió no poder probar su aserto. Más tarde empezó a dudar de que fuese cierto. Hasta ahora, sólo se conocen los cinco números primos 3, 5, 17, 257 y 65.537 dados por la fórmula (ver cap. 25, sec. 4).

Fermat enunció y esbozó [13] la demostración por descenso infinito del siguiente teorema: el área de un triángulo rectángulo cuyos lados sean números racionales no puede ser un cuadrado perfecto. El esbozo de prueba es el único que diera jamás, y se sigue como corolario que la resolución de $x^4 + y^4 = z^4$ en números enteros es imposible.

Con respecto a los números poligonales, Fermat enunció en su

---

[12] *Oeuvres*, 2, 206.

[13] *Oeuvres*, 1, 340; 3, 271.

libro de Diofanto el importante teorema de que todo entero positivo es, o bien triangular, o la suma de dos o tres números triangulares; todo entero positivo es, o bien un cuadrado, o la suma de 2, 3 ó 4 cuadrados; todo entero positivo es, o bien pentagonal, o la suma de 2, 3, 4 ó 5 números pentagonales; y así sucesivamente para números poligonales de órdenes más elevados. Hizo falta mucho trabajo para probar estos resultados, que son correctos sólo si se incluyen 0 y 1 como números poligonales. Fermat afirma haber demostrado estos teoremas por el método del descenso infinito.

Los números perfectos fueron, como sabemos, estudiados por los griegos, y Euclides dio el resultado básico de que $2^{n-1}(2^n - 1)$ es perfecto si $2^n - 1$ es primo. Para $n = 2, 3, 5$ y 7, los valores de $2^n - 1$ son primos, de forma que 6, 28, 496 y 8128 son números perfectos (como ya fue observado por Nicómaco). Un manuscrito de 1456 da correctamente 33.550.336 como el quinto número perfecto; corresponde a $n = 13$. En su *Epitome* (1536), Hudalrich Regius da también este quinto número perfecto. Pietro Antonio Cataldi (1552-1626) observó en 1607 que $2^n - 1$ es compuesto si n lo es, y comprobó que $2^n - 1$ es primo para $n = 13$, 17 y 19. En 1644 Marin Mersenne dio otros valores. Fermat trabajó también en el problema de los números perfectos. Consideró en qué casos es $2^n - 1$ primo, y en carta a Mersenne de junio de 1640 enunció estos teoremas: (a) Si n no es primo, $2^n - 1$ no es primo; (b) si n es primo, $2^n - 1$, de ser divisible, sólo lo es por números primos de la forma $2kn + 1$. En la actualidad se conocen unos veinte números perfectos. Es una cuestión abierta si los hay impares.

Redescubriendo una regla enunciada en primer lugar por Tâbit ibn Qorra, Fermat dio en 1636 un segundo par de números amigos, 17.926 y 18.416 (el primero, 220 y 284, ya lo había dado Pitágoras), y Descartes, en una carta a Mersenne, dio un tercer par, 9.363.548 y 9.437.506.

Fermat volvió a descubrir el problema de resolver $x^2 - Ay^2 = 1$, cuando A es entero no cuadrado. El problema tiene una larga historia entre griegos e hindúes. En una carta de febrero de 1657 a Frénicle, Fermat enunció el teorema de que $x^2 - Ay^2 = 1$ tiene un número ilimitado de soluciones cuando A es positivo y no es un cuadrado perfecto [14]. Euler llamó a ésta «ecuación de Pell», erróneamente,

[14] *Oeuvres*, 2, 333-35.

como ahora sabemos. En la misma carta [15], Fermat desafió a todos los matemáticos a encontrar una infinidad de soluciones enteras. Lord Brouncker dio algunas soluciones, aunque no demostró que hubiera infinitas. Wallis sí resolvió completamente el problema y dio sus soluciones en cartas de 1657 y 1658 [16] y en el capítulo 98 de su *Algebra*. Fermat también afirmó poder probar cuándo $x^2 - Ax^2 = B$, con A y B dados, es soluble, y poder resolverla. No sabemos cómo resolvió Fermat dichas ecuaciones, aunque dijo en una carta de 1658 que había usado el método del descenso para la primera.

8. La relación entre el álgebra y la geometría

El álgebra, como hemos podido ver, se expandió enormemente durante los siglos XVI y XVII. Por haber estado tan ligada a la geometría, antes de 1500 las ecuaciones de grado superior al tercero eran consideradas irreales. Cuando el estudio de las ecuaciones de grado superior se impuso a los matemáticos (por ejemplo, por el uso de las identidades trigonométricas como ayuda en el cálculo de tablas) o fue sugerido como extensión natural de las ecuaciones de tercer grado, la idea pareció absurda a muchos matemáticos. Así, en su edición del *Coss* (Algebra) de Rudolff, Stifel dice: «Ir más allá del cubo como si hubiese más de tres dimensiones... va contra la naturaleza.»

No obstante, el álgebra se impuso sobre las limitaciones del pensamiento geométrico, aunque la relación entre álgebra y geometría siguió siendo complicada. El mayor problema era justificar el razonamiento algebraico, y la solución, durante el siglo XVI y parte del XVII, fue apoyarse en el significado geométrico equivalente de cada desarrollo algebraico. Pacioli, Cardano, Tartaglia, Ferrari y otros dieron pruebas geométricas de las reglas algebraicas. También Vieta estaba profundamente ligado a la geometría: escribe, por ejemplo, $A^3 + 3B^2A = Z^3$, donde A es la incógnita y B y Z son constantes, con objeto de que cada término sea de tercer grado y represente, por tanto, un volumen. Como veremos, sin embargo, la posición de Vieta con respecto al álgebra era de transición. Barrow y Pascal pusieron de hecho objeciones al álgebra, y después a los métodos analíticos de la

---

[15] *Oeuvres*, 2, 333-35; 3, 312-13.

[16] Fermat, *Oeuvres*, 3, 457-80, 490-503.

geometría de coordenadas, así como al cálculo, por carecer el álgebra de justificación.

La dependencia del álgebra de la geometría empezó a invertirse en cierta manera cuando Vieta, y después Descartes, emplearon el álgebra para resolver problemas de construcciones geométricas. La motivación de muchos de los desarrollos algebraicos que aparecen en *In Artem Analyticam Isagoge* de Vieta es la resolución de problemas geométricos y la sistematización de construcciones geométricas. Típico de la aplicación del álgebra a la geometría en la obra de Vieta es el siguiente problema de sus *Zeteticorum Libri Quinque:* dadas el área de un rectángulo y la razón de sus lados, hallar los lados del rectángulo. Toma el área como *B planum* y la razón del lado mayor al menor como S a R. Sea A el lado mayor. Entonces RA/S es el lado menor. Por tanto, *B planum* es igual a (R/S) (A cuadrado). Multiplicando por S obtenemos la ecuación final, $BS = RA^2$. Vieta muestra entonces cómo con esta ecuación se puede construir A con regla y compás a partir de las cantidades conocidas, B y R/S. La idea es que, si se deduce que una longitud deseada x satisface la ecuación $ax^2 + bx + c = 0$, se sabe que

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

y se puede construir x efectuando con a, b y c las construcciones geométricas indicadas en la expresión algebraica del segundo miembro.

El álgebra, para Vieta, significaba un procedimiento especial de descubrimiento; era análisis en el sentido de Platón, que lo oponía a la síntesis. Teón de Alejandría, que introdujo el término «análisis», lo definió como el proceso que comienza con la presunción de lo que se busca y llega por deducción a una verdad conocida. Esta es la razón por la cual Vieta llamó a su álgebra «arte analítico», pues llevaba a cabo el proceso de análisis, particularmente en problemas geométricos. Este fue, de hecho, el punto de partida del pensamiento de Descartes sobre la geometría analítica, y su trabajo en la teoría de ecuaciones estuvo motivado por el deseo de emplearlas mejor en la resolución de problemas geométricos.

La interdependencia entre álgebra y geometría puede verse también en la obra de Marino Ghetaldi (1566-1627), alumno de Vieta. Hizo un estudio sistemático de la solución algebraica de determinados

problemas geométricos en uno de los libros de su *Apollonius Redivivus* (Apolonio redivivo o modernizado, 1607). Recíprocamente, dio pruebas geométricas de reglas algebraicas y construyó geométricamente las raíces de ecuaciones algebraicas. Un trabajo completo sobre este tema es su *De resolutione et compositione mathematica* (1630), publicado póstumamente.

Hallamos también en los siglos XVI y XVII la conciencia de la necesidad de desarrollar el álgebra como sustitución de los métodos geométricos introducidos por los griegos. Vieta vio la posibilidad de utilizar el álgebra para tratar la igualdad y la proporción de magnitudes, sin importar si tales magnitudes surgían de problemas geométricos, físicos o comerciales. De aquí que no dudase en considerar ecuaciones de grado superior y en utilizar el método algebraico, abrigando la idea de una ciencia deductiva de las magnitudes que emplease el simbolismo. Mientras que el álgebra era para Vieta principalmente una vía perfecta hacia la geometría, su visión era lo bastante grandiosa como para percibir en ella una vida y significado propios. Bombelli dio pruebas algebraicas aceptables para su tiempo, sin el uso de la geometría. Stevin afirmaba que lo que se podía hacer en geometría podía hacerse en aritmética y álgebra. El libro de Harriot *Artis Analyticae Praxis* (1631) extendió, sistematizó e hizo emerger algunas de las implicacaciones de la obra de Vieta. El libro es muy parecido a un texto moderno de álgebra. Es más analítico que ninguna obra anterior de álgebra, y presenta grandes avances en el simbolismo; fue ampliamente utilizado.

También Descartes empezó a ver grandes potencialidades en el álgebra, comenzando, según él, donde Vieta terminó. No considera el álgebra como una ciencia en el sentido de proporcionar conocimiento sobre el mundo físico. Ciertamente, afirma, tal conocimiento consta de geometría y mecánica. Ve en el álgebra un poderoso método de guía del razonamiento con cantidades desconocidas y abstractas, sobre todo. En su visión, el álgebra mecaniza la matemática de forma que el pensamiento y los procesos se simplifican y no requieren gran esfuerzo de la mente. La creación matemática podría volverse casi automática.

El álgebra, para Descartes, precede a las demás ramas de la matemática. Es una extensión de la lógica útil para tratar cantidades, y, en este sentido, incluso más fundamental que la geometría, es decir, lógicamente anterior a ella. Por todo ello, buscó un álgebra independiente y sistemática, en vez de una colección sin plan ni fundamento

de símbolos y métodos ligados a la geometría. Hay un esbozo de tratado de álgebra, conocido como *Le Calcul* (1638), escrito por el propio Descartes o bajo su dirección, que considera el álgebra como una ciencia específica. Este álgebra está desprovista de sentido; es una técnica de cálculo, y forma parte de su búsqueda general del método.

La visión de Descartes del álgebra como una extensión de la lógica que trata de la cantidad le sugirió la posibilidad de crear una ciencia algebraica más amplia que considerase, además de la cantidad, otros conceptos y pudiese utilizarse para cualquier tipo de problema. Hasta los principios y métodos lógicos podrían expresarse simbólicamente, y se utilizaría todo el sistema para mecanizar el razonamiento. Descartes llamó a esta noción «matemática universal». La idea es vaga en sus realizaciones, y no llegó a investigarla en profundidad. A pesar de todo, fue el primero en asignar al álgebra un lugar fundamental en el sistema de conocimiento.

Esta visión del álgebra, considerada en su plenitud en primer lugar por Leibniz y finalmente desarrollada en la lógica simbólica (cap.51, sec. 4), fue también la de Barrow, aunque en menor amplitud. Barrow, amigo, profesor y predecesor de Newton en la cátedra Lucasiana de matemáticas en Cambridge, no consideraba el álgebra como parte de la matemática propiamente dicha, sino como una formalización de la lógica. Para él, sólo la geometría era matemática, y la aritmética y el álgebra trataban de magnitudes geométricas expresadas en símbolos.

Con independencia de la filosofía del álgebra de Descartes y Barrow y de las potencialidades que ambos puedan haber visto en ella como ciencia universal del razonamiento, el efecto práctico del uso creciente de técnicas aritméticas y algebraicas fue el de establecer el álgebra como rama de las matemáticas independiente de la geometría. Para la época, Descartes dio un paso importante al escribir a^2 para representar tanto una longitud como un área, cuando Vieta había insistido en que una segunda potencia debía representar un área. Descartes llamó la atención sobre su utilización de a^2 como un número, observando explícitamente que se había apartado de sus predecesores. Dice que x^2 es una cantidad tal que $x^2{:}x = x{:}1$. Igualmente, dice en *La Géometrie* que un producto de líneas puede ser una línea; pensaba en las cantidades, y no en términos geométricos como habían hecho los griegos. Le parecía perfectamente claro que el cálculo algebraico era independiente de la geometría.

John Wallis, influido por Vieta, Descartes, Fermat y Harriot, fue

mucho más lejos que ellos al liberar la aritmética y el álgebra de la representación geométrica. En su *Algebra* (1685) dedujo algebraicamente todos los resultados del libro V de Euclides. Abandonó también la limitación de homogeneidad en las ecuaciones algebraicas en *x* e *y*, concepto que se había mantenido debido a que dichas ecuaciones se derivaban de problemas geométricos. Veía en el álgebra brevedad y transparencia.

Aunque a Newton le entusiasmaba la geometría, hallamos por vez primera en su *Arithmetica Universalis* una afirmación de la importancia básica de la aritmética y el álgebra en oposición a la geometría. Descartes y Barrow todavía favorecían a la geometría como rama fundamental de las matemáticas. Newton necesitó y utilizó el lenguaje algebraico para el desarrollo del cálculo, cuya manipulación óptima era la algebraica. Y en lo que se refiere a la supremacía del álgebra sobre la geometría, las necesidades del cálculo diferencial e integral iban a ser decisivas.

Así, hacia 1700, el álgebra había alcanzado una situación que le permitía mantenerse sobre sus propios pies. La única dificultad era que no tenía lugar donde ponerlos. Desde los tiempos de egipcios y babilonios, la intuición, el ensayo y error habían proporcionado algunas reglas de cálculo, la reinterpretación del álgebra geométrica griega había dado otras, y el trabajo algebraico independiente en los siglos XVI y XVII, guiado en parte por la interpretación geométrica, había conducido a muchos resultados nuevos. Pero para el álgebra no había fundamentos lógicos análogos a los que Euclides había dado a la geometría. Esta despreocupación, aparte de las objeciones de Pascal y Barrow, resulta sorprendente, ya que, para esta época, había en Europa plena conciencia de las exigencias de una matemática deductiva rigurosa.

¿Cómo sabían los matemáticos lo que era correcto? Las propiedades de los enteros positivos y las fracciones se deducen de manera tan inmediata de la experiencia con colecciones de objetos que parecen evidentes. Hasta Euclides dejó sin base lógica a los libros de los *Elementos* que tratan de la teoría de números. A medida que se fueron añadiendo nuevos tipos de números al sistema numérico, las reglas de las operaciones ya aceptadas para los enteros positivos y las fracciones se aplicaron a los nuevos elementos, con el pensamiento geométrico como útil guía. Las letras, cuando fueron introducidas, eran simples representaciones de números y podían ser tratadas como tales. Las técnicas algebraicas más complicadas se justificaban bien por argu-

mentos geométricos como los de Cardano, bien por pura inducción sobre casos específicos; pero ninguno de estos procedimientos era lógicamente satisfactorio. Ni siquiera la invocación a la geometría daba la base lógica para los números negativos, irracionales o complejos, ni justificaba, por ejemplo, la argumentación de que si un polinomio es negativo para $x = a$ y positivo para $x = b$, debe anularse entre a y b.

No obstante todo ello, los matemáticos procedieron alegre y confiadamente al empleo de la nueva álgebra. Wallis, de hecho, afirmaba que los métodos del álgebra no eran menos legítimos que los de la geometría. Sin darse cuenta, los matemáticos estaban a punto de entrar en una nueva era en la que inducción, intuición, ensayo y error y argumentos físicos iban a servir de bases para las demostraciones. El problema de construir un fundamento lógico para el sistema numérico y el álgebra era difícil, mucho más de lo que ningún matemático del siglo XVII podía haber percibido. Y es una suerte que los matemáticos fuesen crédulos, y hasta ingenuos, y no lógicamente escrupulosos; pues la creación libre debe preceder a la formalización y la fundamentación lógica, y el más grande período de creación matemática acababa de comenzar.

Bibliografía

Ball, W. W. R.: *A Short Account of the History of Mathematics*. Dover (reprint), 1960, cap. 12.

Boyer, Carl B.: *History of Analytic Geometry*. Scripta Mathematica, 1956, cap. 4.

—: *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Cajori, Florian: *Oughtred, A Great Seventeenth Century Teacher of Mathematics*, Open Court, 1916.

—: *A History of Mathematics*, 2nd. ed., Macmillan, 1919, pp. 130-59.

—: *A History of Mathematical Notations*, Open Court, 1928, vol. 1.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1900, Johnson Reprint Corp., 1965, vol. 2, pp. 369-806.

Cardan, G.: *Opera Omnia*, 10 vols., 1663, Johnson Reprint Corp., 1964.

Cardano, Girolamo: *The Great Art*, trad. T. R. Witmer, Massachusetts Institute of Technology Press, 1968.

Coolidge, Julian L.: *The Mathematics of Great Amateurs*. Dover (reimpresión), 1963, caps. 6-7.

David, F. N.: *Games, Gods and Gambling: The Origins and History of Probability*, Hafner, 1962.
Descartes, René: *The Geometry*, Dover (reimpresión), 1954, Libro 3.
—: *Discurso del método*, Madrid, Alfaguara, 1987.
Dickson, Leonard E.: *History of the Theory of Numbers*, Carnegie Institution, 1919-23, Chelsea (reimpresión), 1951, vol. 2, cap. 26.
Fermat, Pierre de: *Oeuvres*, 4 vols. y suplemento, Gauthier-Villars, 1891-1912, 1922.
Heath, Sir Thomas L.: *Diophantus of Alexandria*, 2 ed., Cambridge University Press, 1910. Dover (reimpresión), 1964, Suplemento, secs. 1-5.
Hobson, E. W.: *John Napier and the Invention of Logarithms*. Cambridge University Press, 1914.
Klein, Jacob: *Greek Mathematical Thought and the Origin of Algebra*. Massachusetts Institute of Technology Press, 1968. Contiene una traducción del *Isagoge* de Vieta.
Knott, C. G.: *Napier Tercentenary Memorial Volume*, Longmans Green, 1915.
Montucla, J. F.: *Histoire des mathématiques*, Albert Blanchard (reimpresión), 1960, vol. 1, parte, 3, libro 3; vol. 2, parte 4, libro 1.
Mordell, J. L.: *Three Lectures on Fermat's Last Theorem*, Cambridge University Press, 1921.
Morley, Henry: *Life of Cardan*, 2 vols., Chapman and Hall, 1854.
Newton, Sir Isaac: *Mathematical Works*, vol. 2, ed. D. T. Whiteside, Johnson Reprint Corp., 1967. Este volumen contiene una traducción de la *Arithmetica Universalis*.
Ore, Oystein: *Cardano: The Gambling Scholar*, Princeton University Press, 1953.
—: «Pascal and the Invention of Probability Theory», *Amer. Math. Monthly*, 67, 1960, 409-19.
Pascal, Blaise: *Oeuvres complètes*. Hachette, 1909.
—: *Obras. 1*, Madrid, Alfaguara, 1981.
Sarton, George: *Six Wings: Men of Science in the Renaissance*, Indiana University Press, 1957, «Second Wing» [Hay ed. cast.: *Seis alas*, Buenos Aires, 1960].
Sarton, George: *Six Wings: Men of Science in the Renaissance*, Indiana University Press, 1957, «Second Wing».
Schneider, I.: «Der Mathematiker Abraham de Moivre (1667-1754)», Archive for History of Exact Sciences, 5, 1968, 177-317.
Scott, J. F.: *A History of Mathematics*, Taylor and Francis, 1958, caps. 6 y 9.
—: *The Scientific Work of René Descartes*, Taylor and Francis, 1952, cap. 9.
—: *The Mathematical Work of John Wallis*, Oxford University Press, 1938.
Smith, David Eugene: *History of Mathematics*, Dover (reimpresión), 1958, vol. 1, caps. 8-9; vol. 2, caps. 4 y 6.
—: *A Source Book in Mathematics*, 2 vols., Dover (reimpresión), 1959.

Struik, D. J.: *A Source Book in Mathematics, 1200-1800*, Harvard University Press, 1969, caps. 1-2.
Todhunter, I.: *A History of the Mathematical Theory of Probability*, Chelsea (reimpresión), 1949, caps. 1-7.
Turnbull, H. W.: *James Gregory Tercentenary Memorial Volume*, Bell and Son, 1939.
—: *The Mathematical Discoveries of Newton*, Blackie and Son, 1945.
Vieta, F.: *Opera Mathematica*, (1646), Georg Olms (reimpresión), 1970.

Capítulo 14
LOS COMIENZOS DE LA GEOMETRIA PROYECTIVA

> Confieso francamente que nunca he sentido gusto por el estudio o la investigación en física o en geometría, a no ser que pudieran servir como medio de llegar a algún tipo de conocimiento de las causas próximas... para el bien y la comodidad de la vida, el mantenimiento de la salud, la práctica de algún arte... pues he observado que una buena parte de las artes se basa en la geometría, como el de tallar la piedra en arquitectura, el de los relojes de sol, y el de la perspectiva en particular.
>
> GIRARD DESARGUES

1. El renacer de la geometría

El resurgir de la actividad creadora en geometría anduvo a la zaga del renacer del álgebra. Aparte de la creación del sistema matemático de perspectiva y de las incidentales aportaciones geométricas de los artistas de Renacimeinto (cap. 12, sec. 2), muy poco de importancia se hizo en geometría desde el tiempo de Pappus hasta 1600. Provocó cierto interés la aparición de numerosas ediciones impresas de las *Secciones Cónicas* de Apolonio, especialmente la notable traducción latina de Federigo Commandino (1509-75) de los libros I al IV en 1566. Los libros V al VII vieron la luz gracias a otros traductores, y una serie de matemáticos, entre ellos Vieta, Willebrord Snell (1580-1626) y Ghetaldi emprendieron la reconstrucción del perdido libro VIII.

Lo que hacía falta para dirigir las mentes de los matemáticos por nuevos caminos eran problemas nuevos. Uno de ellos había sido ya planteado por Alberti: ¿qué propiedades geométricas tienen en común dos secciones de la misma proyección de una figura real? Un gran número de problemas procedía de las ciencias y las necesidades prácticas. El uso que hizo Kepler de las secciones cónicas en su obra

de 1609 dio un enorme impulso al reexamen de dichas curvas y a la búsqueda de propiedades que fueran útiles para la astronomía. La óptica, que venía interesando a los matemáticos desde la época griega, recibió gran atención tras la invención del telescopio y el microscopio, a principios del siglo XVII. El diseño de lentes para estos instrumentos se convirtió en un problema de primera magnitud, que se tradujo en interés por la forma de las superficies, o la de sus curvas generatrices si aquéllas eran de revolución. Las exploraciones geográficas habían creado la necesidad de disponer de mapas y de estudiar las trayectorias de los barcos representados en la esfera y en el mapa. La introdución de la noción de una Tierra móvil exigía nuevos principios en la mecánica que dieran cuenta de las trayectorias de los objetivos móviles, lo que también significaba estudiar curvas. Entre los objetivos móviles, los proyectiles eran los más importantes, ya que los cañones de la época podían disparar ya balas a cientos de metros, siendo, pues, vital predecir su trayectoria y su alcance. El problema práctico de hallar áreas y volúmenes empezó a atraer mayor atención. La *Nova Stereometria Doliorum Vinariorum* (Nueva Ciencia de Medir Volúmenes de Toneles de Vino, 1615) de Kepler inició un nuevo estallido de actividad en este área.

Otra clase de problemas entró en escena como consecuencia de la asimilación de las obras griegas. Los matemáticos empezaron a darse cuenta de que a los métodos griegos de demostración les faltaba generalidad. Hacía falta crear un método especial casi para cada teorema. Esta observación ya la había hecho Agrippa von Nettesheim en 1527, así como Maurolico, traductor de obras griegas y autor de libros sobre secciones cónicas y otras cuestiones matemáticas.

La mayoría de las soluciones dadas a estos nuevos problemas resultaron no ser más que variaciones menores sobre temas antiguos. La forma de tratar las secciones cónicas se vio alterada. Tales curvas empezaron a definirse como lugares geométricos en el plano, en vez de como secciones de un cono, como en la obra de Apolonio. Guidobaldo del Monte, por ejemplo, definió la elipse en 1579 como el lugar de los puntos cuya suma de distancias a los focos es constante. No sólo las cónicas, sino otras curvas griegas más antiguas, como la concoide de Nicomedes, la cisoide de Diocles, la espiral de Arquímedes y la cuadratriz de Hipias fueron estudiadas de nuevo. Se introdujeron nuevas curvas, muy especialmente la cicloide (ver cap. 17, sec. 2). Pero aunque todo este trabajo era útil para difundir las contribuciones de los griegos, nada en él ofrecía novedades en resultados o en

métodos de demostración. La primera innovación trascendente tuvo lugar como respuesta a los problemas planteados por los pintores.

2. Problemas planteados por los trabajos sobre la perspectiva

La idea básica del sistema de perspectiva focal creado por los pintores es el principio de proyección y sección (cap. 12, sec. 1). Una escena real es vista por el ojo, considerado como un punto. Las líneas de fuga desde varios puntos de la escena hacia el ojo se dice que constituyen una proyección. Según dicho principio, la propia pintura debe contener una sección de tal proyección, formada por lo que aparecería en un plano que pasase a través de la proyección.

Supongamos ahora que un ojo situado en *O* (fig. 14.1) mira hacia el rectángulo horizontal *ABCD*. Las líneas trazadas desde *O* hasta los puntos de los cuatro lados del rectángulo constituyen una proyección, de la cual *OA*, *OB*, *OC* y *OD* son líneas típicas. Si ahora interponemos un plano entre el ojo y el rectángulo, las líneas de la proyección lo atravesarán y dejarán marcado en él un cuadrilátero *A'B'C'D'*. Como la sección, es decir, *A'B'C'D'*, crea la misma impresión en el

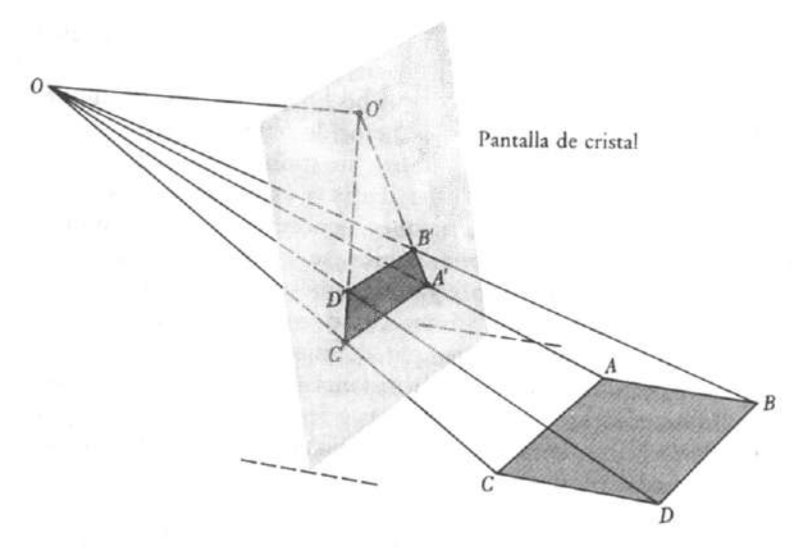


FIGURA 14.1

ojo que el rectángulo original, es razonable preguntarse, como Alberti, qué propiedades geométricas tienen en común la sección y el rectángulo original. Es intuitivamente claro que la figura original y la sección no son ni congruentes ni semejantes, ni tienen la misma área. De hecho, la sección no es necesariamente un rectángulo.

Este problema admite una extensión: consideremos dos secciones distintas de esta misma proyección correspondientes a dos planos distintos que cortan a la proyección en ángulos cualesquiera. ¿Qué propiedades geométricas tienen en común ambas secciones?

El problema puede extenderse aún más. Supongamos que un rectángulo *ABCD* es visto desde dos lugares distintos, O' y O'' (fig. 14.2). Hay entonces dos proyecciones, una determinada por O' y el rectángulo y la otra determinada por O'' y el mismo rectángulo. Si se realiza una sección de cada proyección, visto que cada una de ellas debería tener ciertas propiedades en común con el rectángulo, también ellas deberían tener propiedades geométricas comunes entre sí.

Algunos de los geómetras del siglo XVII se propusieron responder a estas cuestiones. Siempre consideraron los métodos que utilizaban y los resultados que obtenían como parte de la geometría euclídea, y, sin embargo, esos métodos y resultados, aun contribuyendo sin duda en gran medida a ese campo, resultaron ser los comienzos de una

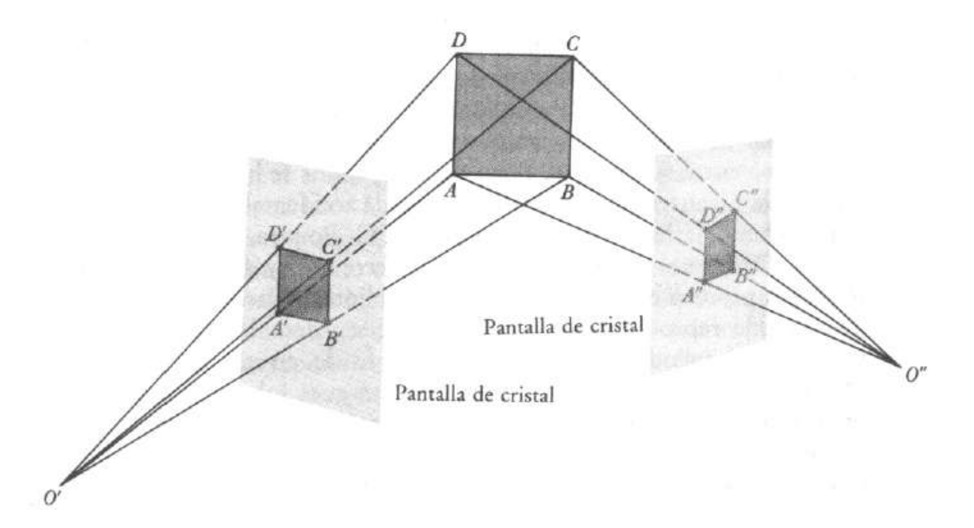


FIGURA 14.2

nueva rama de la geometría, conocida a partir del siglo XIX como «geometría proyectiva». Nos referiremos a las investigaciones en este área con tal nombre, aunque en el siglo XVII no se hacía distinción entre las geometrías euclídea y proyectiva.

3. La obra de Desargues

Quien primero atacó directamente los problemas arriba mencionados fue el autodidacta Girard Desargues (1591-1661), que llegó a ser oficial del ejército y después ingeniero y arquitecto. Desargues conocía la obra de Apolonio y pensó en la posibilidad de introducir nuevos métodos de demostración de teoremas sobre cónicas, cosa que hizo, y con plena conciencia de la potencia de tales métodos. Desargues quería también mejorar la formación y las técnicas de los artistas, ingenieros y talladores de piedra; la teoría por sí misma le interesaba poco. Empezó por poner en orden muchos teoremas utilizables, difundiendo inicialmente sus resultados en cartas y hojas impresas. Dio también clases gratis en París. Después escribió varios libros, uno de ellos de enseñanza de canto para niños, y otro de aplicación de la geometría a la albañilería y al tallado de piedras.

Su texto principal, precedido en 1636 por un folleto sobre perspectiva, fue *Brouillon project d'une atteinte aux événements des recontres du cône avec un plan* (Borrador de un ensayo de tratado de los resultados de los encuentros de un cono con un plano, 1939) [1]. Este libro trata de lo que ahora llamaríamos métodos proyectivos en geometría. Desargues imprimió unos cincuenta ejemplares y los distribuyó entre sus amigos. Poco después, todos se habían perdido. Una copia manuscrita de La Hire fue hallada accidentalmente en 1845 por Michel Chasles y reproducida por N. G. Poudra, quien editó la obra de Desargues en 1864. Finalmente, Pierre Moisy descubrió hacia 1950 una primera edición de 1639 en la Biblioteca Nacional de París, que ha sido reproducida. Este ejemplar recién descubierto contiene también un apéndice y una importante fe de erratas del propio Desargues. El teorema fundamental de Desargues sobre triángulos y otros resultados fueron publicados en 1648 en un apéndice de un libro sobre perspectiva de un amigo suyo, Abraham Bosse (1602-76). En este libro, *Manière universelle de M. Desargues, pour pratiquer la*

---

[1] *Oeuvres*, 1, 103-230.

perspective, Bosse trató de popularizar los métodos prácticos de Desargues.

Desargues empleaba una curiosa terminología, parte de la cual ya aparecía en la obra de Alberti. Llamaba «palmas» a las rectas. Cuando ciertos puntos aparecían señalados en la recta, lo llamaba «tronco». Sin embargo, una recta con tres pares de puntos en involución (ver más abajo) era un «árbol». La intención de Desargues al introducir esta nueva terminología era ganar en claridad al evitar las ambigüedades presentes en los términos más usuales. Pero su lenguaje y sus extrañas ideas hacían difícil la lectura del libro. Mersenne, Descartes, Pascal y Fermat lo llamaron loco. Incluso Descartes, al enterarse de que Desargues había introducido un nuevo método para tratar las cónicas, escribió a Mersenne que no veía cómo nadie pudiera hacer nada nuevo en cónicas que no fuese con ayuda del álgebra. Sin embargo, al conocer Descartes los detalles del trabajo de Desargues, lo respetó profundamente. Fermat consideraba a Desargues el verdadero fundador de la teoría de las secciones cónicas, y hallaba su libro rico en ideas, que al parecer poseía. Pero la falta general de aprecio hacia su obra repugnó a Desargues, que se retiró a su hacienda.

Antes de dar cuenta de los teoremas de Desargues, hemos de introducir un nuevo convenio sobre paralelas. Alberti había apuntado que dos rectas que son paralelas en una escena real deben ser dibujadas de forma que se corten en un punto, a no ser que sean paralelas al lienzo. Así, las líneas $A'B'$ y $C'D'$ de la figura 14.1 más arriba, que se corresponden con las paralelas AB y CD, según el principio de proyección y sección deben cortarse en algún punto O'. De hecho, O y AB determinan un plano, así como O y CD. Estos dos planos cortan al cristal en $A'B'$ y $C'D'$, pero como se cortan en O, esos dos planos han de tener una recta común; esta recta corta al cristal en cierto punto O', que es también la intersección de $A'B'$ y $C'D'$. El punto O' no corresponde a ningún punto ordinario de AB ni de CD. De hecho, la recta OO' es horizontal, y por tanto paralela a AB y CD. El punto O' se llama punto impropio o de anulación, pues no tiene correspondiente en las rectas AB y CD, mientras que cualquier otro punto de $A'B'$ y $C'D'$ se corresponde con un punto definido de AB y CD, respectivamente.

Para completar la correspondencia entre los puntos de las líneas $A'B'$ y AB, así como entre los puntos de $C'D'$ y CD, Desargues introdujo un nuevo punto en AB y CD, llamado punto del infinito, que se añade a los puntos usuales de ambas líneas y debe considerarse

como su punto común. Además, cualquier paralela a *AB* o *CD* debe contenerlo y cortar a *AB* y *CD* en él. Cualquier conjunto de rectas paralelas con dirección distinta de *AB* o *CD* tiene asimismo un punto del infinito. Como cada conjunto de rectas paralelas tiene un punto en común y existe un número infinito de conjuntos distintos de paralelas, el convenio de Desargues introduce una infinidad de nuevos puntos en el plano euclídeo. Hizo la hipótesis adicional de que éstos yacen en una recta, que corresponde a la línea del horizonte en la sección. Se añade así una nueva línea recta a las ya existentes en el plano euclídeo. Se supone que un conjunto de planos paralelos tiene en común la recta del infinito de cada uno de ellos; es decir, todos los planos paralelos se cortan en una recta.

La adición de un nuevo punto a cada recta no contradice axioma o teorema alguno de la geometría euclídea, aunque requiere ciertos cambios de terminología. Las rectas no paralelas siguen cortándose en puntos ordinarios, mientras que dos paralelas se cortan en el «punto del infinito» de cada una. Este convenio sobre los puntos del infinito en geometría euclídea es cómodo, pues evita considerar casos particulares. Por ejemplo, ahora se puede decir que dos rectas cualesquiera se cortan exactamente en un punto. Pronto veremos otras ventajas de este convenio.

También Kepler, en 1604, había añadido un punto del infinito a las rectas paralelas, pero por razones distintas. Cada recta que pasa por *P* (fig. 14.3) y corta a *l* tiene un punto común, *Q*, con ésta. Sin embargo, no corresponde ningún punto a *PR*, la paralela a *l* por *P*. Al añadir un punto en el infinito común a *PR* y *l*, Kepler podía afirmar que cada recta que pasa por *P* corta a *l*. Además, una vez que *Q* se ha ido «al infinito» por la derecha y *PQ* se ha convertido en *PR*, el punto de intersección de *PR* y *l* puede considerarse como un punto en el infinito a la izquierda de *P*; y a medida que *PR* sigue girando alrededor de *P*, el punto *Q'* de intersección de *PR* y *l* se moverá hacia la izquierda. Así, se mantiene la continuidad de la intersección de *PR*

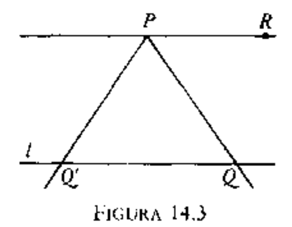


FIGURA 14.3

y *l*. En otras palabras, Kepler (y Desargues) consideraba que los dos «extremos» de la recta se cortan «en el infinito», de forma que la recta tiene la estructura de un círculo. De hecho, Kepler pensaba en una recta como un círculo con su centro en el infinito.

Una vez introducidos sus puntos y rectas del infinito, Desargues pudo enunciar un teorema básico, llamado hoy «teorema de Desargues». Consideremos (fig. 14.4) el punto O y el triángulo ABC. Las rectas que parten de O hacia los diversos puntos de los lados del triángulo constituyen, como sabemos, una proyección. Una sección de esta proyección contendrá, pues, un triángulo $A'B'C'$, donde A' corresponde a A, B' a B y C' a C. Ambos triángulos, ABC y $A'B'C'$, se dicen perspectivos desde el punto O. El teorema de Desargues afirma entonces: los pares de lados homólogos AB y $A'B'$, BC y $B'C'$, y AC y $A'C'$ (o sus prolongaciones) de dos triángulos perspectivos desde un punto se cortan en tres puntos alineados. Recíprocamente, si los tres pares de lados homólogos de dos triángulos se cortan en tres puntos alineados, entonces las rectas que unen los vértices homólogos se cortan en un punto. Refiriéndonos específicamente a la figura, el teorema propiamente dicho afirma que, dado que AA', BB' y CC' se cortan en un punto O, los lados AC y $A'C'$ se cortan en un punto P, AB y $A'B'$ se cortan en un punto Q y BC y $B'C'$ se cortan en un punto R, y P, Q y R se hallan sobre una recta.

Aunque el teorema es cierto tanto si los triángulos están en el

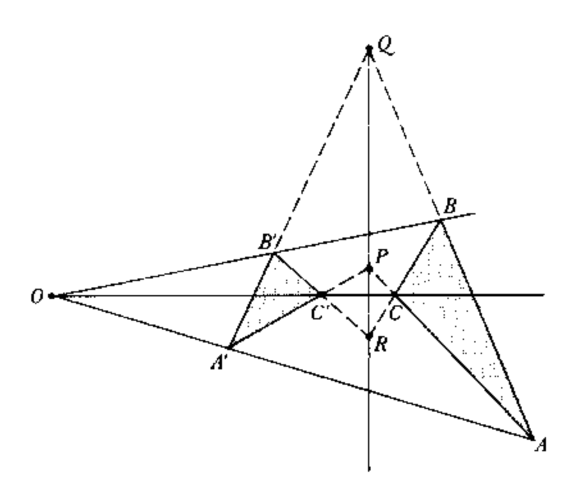


FIGURA 14.4

mismo o en diferentes planos, la demostración sólo es sencilla en el último caso. Desargues probó el teorema y su recíproco, tanto para el caso bidimensional como el tridimensional.

En el apéndice de la obra de Bosse de 1648 aparece otro fundamental resultado de Desargues: la invariancia de la razón doble por proyecciones. La razón doble de los segmentos formados por cuatro puntos A, B, C y D de una recta (fig. 14.5) es, por definición, $\frac{BA}{BC} : \frac{DA}{DC}$. Pappus ya había introducido esta razón (cap. 5, sec. 7) y había probado que es la misma en las dos rectas AD y $A'D'$. Menelao también probó un teorema semejante sobre círculos máximos en la esfera (cap. 5, sec. 6). Ninguno de ellos, sin embargo, pensaba en términos de proyecciones y secciones. Desargues sí, y demostró que la razón doble es la misma en cualquier sección de una proyección.

En su obra fundamental de 1639 trata del concepto de involución, también introducido por Pappus, y al que Desargues dio nombre. Se dice que dos pares de puntos A, B y A', B' constituyen una involución si existe un punto especial O, llamado centro de la involución, sobre la recta que contiene a los cuatro puntos, tal que $OA \cdot OB = OA' \cdot OB'$ (fig. 14.6). Análogamente, se dice que tres pares de puntos A, B, A', B' y A'', B'' constituyen una involución si $OA \cdot OB = OA' \cdot OB' = OA'' \cdot OB''$. Los puntos A y B, A' y B' y A'' y B'' se dice que son conjugados. Si existe un punto E tal que $OA \cdot OB = \overline{OE}^2$, E se llama punto doble. En tal caso, existe un segundo punto doble F, y O es el punto medio de EF. El conjugado de O es el punto del infinito. Desargues empleó el teorema de Menelao sobre la recta que corta los lados de un triángulo para demostrar que si los cuatro puntos A, B, A' y B' forman una involución (fig. 14.7), y si se proyectan desde P en los puntos A_1, B_1, A_1' y B_1' de otra recta, entonces el segundo conjunto de puntos forma también una involución.

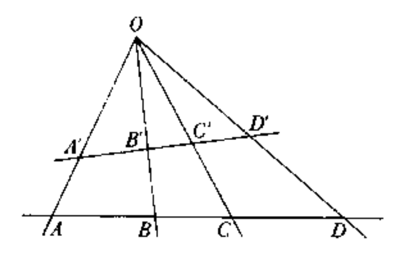


FIGURA 14.5

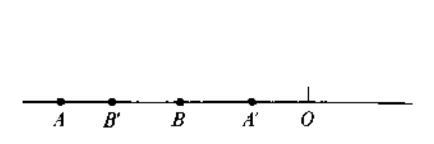


FIGURA 14.6

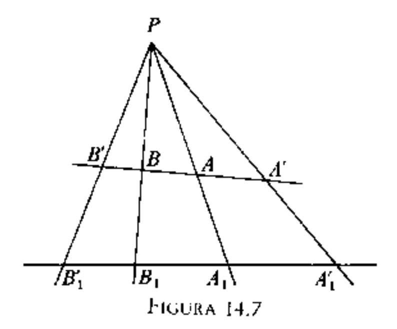


FIGURA 14.7

Desargues demostró un teorema fundamental sobre involuciones. Para explicarlo, consideremos primero el concepto de cuadrilátero completo, noción ya tratada en parte por Pappus. Sean *B*, *C*, *D* y *E* cuatro puntos cualesquiera de un plano (fig. 14.8), de los cuales no haya tres alineados. Estos determinan seis rectas, que son los lados del cuadrilátero completo. Dos lados son opuestos cuando no tienen en común ninguno de los cuatro puntos, de forma que *BC* y *DE* son opuestos, como *CD* y *BE* y también *BD* y *CE*. Las intersecciones *O*, *F* y *A* de los tres pares de lados opuestos son los puntos diagonales del cuadrilátero. Supongamos ahora que los cuatro vértices *B*, *C*, *D* y *E* están en un círculo, y supongamos que una recta *PM* corta a los pares de lados en *P*, *Q*; *I*, *K*; y *G*, *H*; y al círculo en el par *L*, *M*. Estos cuatro pares de puntos son pares de una involución.

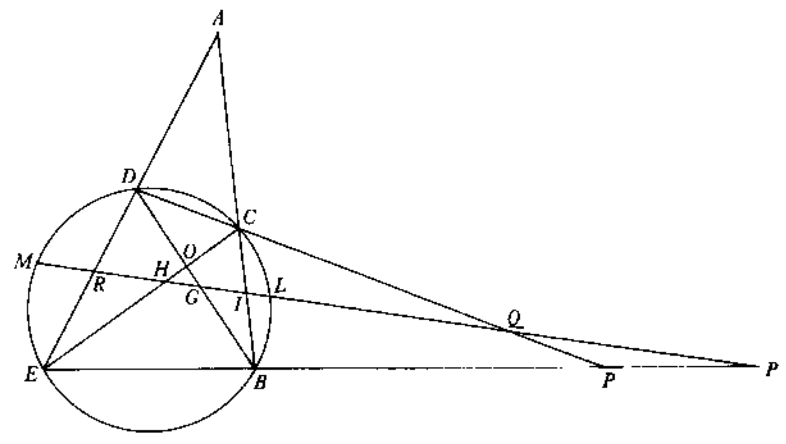


FIGURA 14.8

Supongamos además que la figura completa se proyecta desde un punto exterior al plano de la figura, y se efectúa una sección de esta proyección. El círculo dará origen en la sección a una cónica, y cada recta de la figura original dará lugar a su vez a una recta. En particular, el cuadrilátero inscrito en el círculo dará lugar a un cuadrilátero inscrito en la cónica. Como una involución se proyecta sobre una involución, se sigue el siguiente resultado, importante y general: si un cuadrilátero se halla inscrito en una cónica, cualquier recta que no pase por un vértice corta a la cónica y a los pares de lados opuestos del cuadrilátero completo en cuatro pares de puntos de una involución.

Desargues introduce a continuación la noción de cuaterna armónica. Los puntos A, B, E y F forman una cuaterna armónica si A y B son un par de puntos conjugados con respecto a los puntos dobles E y F de una involución. (La definición actual de que la razón doble de la cuaterna armónica sea igual a -1 es posterior.)[2] Como una involución se proyecta en una involución, una cuaterna armónica se proyecta en una cuaterna armónica. Desargues muestra entonces que si un miembro de una cuaterna armónica es el punto del infinito (de la recta de los cuatro), el otro punto del par es el punto medio del segmento que une los otros dos puntos de la cuaterna. Además, si A, B, A' y B' forman una cuaterna armónica (fig. 14.9), y se considera la proyección desde O, entonces OA' es perpendicular a OB', OA' es la bisectriz del ángulo AOB y OB' es la bisectriz del ángulo suplementario.

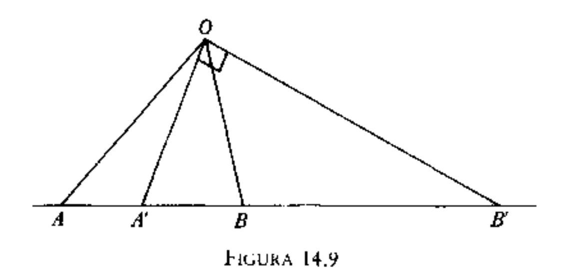


FIGURA 14.9

Con la noción de cuaterna armónica a mano, Desargues sigue con la teoría de polos y polares, ya introducida por Apolonio. Desargues parte de un círculo y un punto exterior A (fig. 14.10). Entonces, sobre cada recta por A que corte al círculo en C y D, por ejemplo, hay un

[2] Se debe a Möbius: *Barycentrysche Calcul* (1827), p. 269.

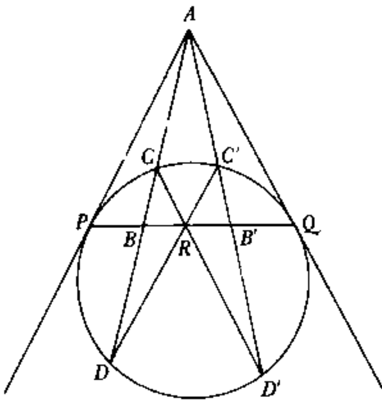


FIGURA 14.10

cuarto armónico B. Para todas esas rectas, el cuarto armónico yace en una recta, la polar del punto A. Así, la recta BB' es la polar de A. Supongamos, además, que introducimos un cuadrilátero completo del cual A sea un punto diagonal, y cuyos vértices estén en el círculo. Sean C, D, D' y C' estos vértices. Entonces, la polar de A pasa por los otros dos puntos diagonales del cuadrilátero completo. (En la figura, R es uno de los puntos diagonales.) Lo mismo se puede decir cuando el punto A se halla en el interior del círculo. Si el punto A es exterior al círculo, la polar de A une los puntos de contacto de las tangentes desde A al círculo (los puntos P y Q de la figura).

Tras demostrar los asertos anteriores para el círculo, Desargues usa una vez más la proyección desde un punto exterior al plano de la figura y una sección de dicha proyección para probar que los resultados se verifican para una cónica cualquiera.

Desargues define un diámetro de una cónica como la polar de un punto del infinito. Vamos a ver que esta definición coincide con la de Apolonio. Consideremos una familia de paralelas que cortan a la cónica (fig. 14.11). Estas rectas tienen en común un punto del infinito. Si $A'B'$ es una de ellas, el conjugado armónico con respecto a A' y B' del punto del infinito de $A'B'$ es el punto medio B de la cuerda $A'B'$. Análogamente, B_1, conjugado armónico del punto del infinito con respecto a A'_1 y B'_1, es el punto medio de la cuerda $A'_1B'_1$. Estos puntos medios de la familia de cuerdas paralelas están en una línea recta, que es un diámetro según la definición de Apolonio. Desargues prueba

entonces una serie de hechos sobre diámetros, diámetros conjugados y asíntotas a hipérbolas.

Vemos, pues, que Desargues no sólo introdujo nuevos conceptos, en particular los elementos del infinito, obteniendo muchos resultados, sino que, sobre todo, introdujo la proyección y sección como nuevo método de demostración, unificando así el estudio de los distintos tipos de cónicas, que Apolonio había tratado por separado. Desargues fue uno de los matemáticos más originales de un siglo rico en genios.

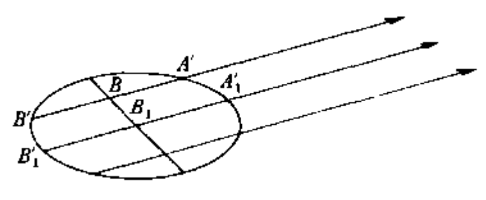


FIGURA 14.11

4. La obra de Pascal y La Hire

La segunda contribución de importancia a la geometría proyectiva fue la de Blaise Pascal (1623-62). Nacido en Clermont, Francia, fue un niño enfermizo y tuvo mala salud a lo largo de su corta vida. Su padre, Etienne Pascal, trató de mantener a su hijo alejado de las matemáticas hasta que tuviese quince o dieciséis años, por considerar que no se debe iniciar a un niño en una materia mientras no tenga edad suficiente para absorberla. Sin embargo, a los doce años de edad, Blaise insistió en saber qué era la geometría, y, una vez que se le dijo, empezó a trabajar en ella por su cuenta.

La familia se había trasladado a París cuando Pascal tenía ocho años. Incluso de niño iba con su padre a las reuniones semanales de la «Académie Mersenne», que más tarde se convertiría en la Academia Libre y, en 1666, en la Academia de Ciencias. Entre sus miembros se hallaban el P. Mersenne, Desargues, Roberval (profesor de matemáticas en el Collège de France), Claude Mydorge (1585-1647) y Fermat.

Pascal dedicó tiempo y energía considerables a la geometría proyectiva. Fue uno de los fundadores del cálculo, e influyó sobre Leibniz en esta cuestión. Como ya hemos visto, también tuvo partici-

pación en los comienzos de la investigación sobre la probabilidad. A los diecinueve años inventó la primera máquina calculadora para ayudar a su padre en su trabajo como tasador de impuestos. Sus aportaciones también incluyen la física, con su trabajo experimental en un original mecanismo para crear el vacío, en el estudio del decrecimiento del peso del aire con la altura y en la aclaración del concepto de presión en los líquidos. La originalidad de su trabajo en física ha sido cuestionada; de hecho, algunos historiadores de la ciencia lo han descrito como divulgativo o como plagiario, según lo compasivos que quisieran ser.

Pascal fue grande en muchos otros campos. Fue un maestro de la prosa francesa; sus *Pensées* y sus *Lettres provinciales* son clásicos de la literatura. Fue también polemista famoso en teología. Desde la infancia trató de reconciliar la fe religiosa con el racionalismo de las matemáticas y la ciencia, y ambos compitieron por su energía y su tiempo a lo largo de su vida. Pascal, como Descartes, creía que las verdades de la ciencia deben apelar claramente a los sentidos o a la razón, o ser consecuencias lógicas de tales verdades. No hallaba lugar para conjurar misterios en materias de ciencia y matemáticas. «Nada que tenga que ver con la fe puede ser objeto de la razón.» En materia de ciencia, en la cual sólo interviene nuestro pensamiento natural, la autoridad es inútil; la razón por sí misma tiene bases para tal conocimiento. Sin embargo, los misterios de la fe se ocultan a los sentidos y a la razón, y deben ser aceptados por la autoridad de la Biblia. Condena a los que usan la autoridad en la ciencia o la razón en la teología, considerando, no obstante, el nivel de la fe superior al de la razón.

La religión dominó sus pensamientos a partir de los veinticuatro años, aunque siguió realizando un trabajo matemático y científico. Creía que el afán por la ciencia por mero disfrute era incorrecto. Hacer del disfrute el fin principal de la investigación era corromperla, porque uno llegaba a adquirir «una codicia o lujuria por aprender, un apetito disoluto por el conocimiento... Un estudio así de la ciencia surge de un interés principal por uno mismo como el centro de las cosas, en lugar de preocuparse de buscar fuera, entre todos los fenómenos naturales que nos rodean, la presencia de Dios y Su gloria».

En su obra matemática fue mayormente intuitivo; anticipó grandes resultados, hizo soberbias conjeturas y supo ver métodos más directos. En etapas más avanzadas de su vida llegó a considerar la

intuición como la fuente de toda verdad. Varias de sus frases sobre este tema se han hecho famosas: «El corazón tiene sus propias razones, que la razón desconoce.» «La razón es el lento y tortuoso método por el cual los que no conocen la verdad la descubren.» «Humíllate, razón impotente.»

Si hemos de guiarnos por la carta que escribió a Fermat el 10 de agosto de 1660, Pascal parece haberse vuelto en cierta medida contra las matemáticas hacia el fin de su vida: «Hablando francamente de las matemáticas, las encuentro el ejercicio más elevado del espíritu; pero al mismo tiempo sé que es tan inútil que hago poca distinción entre un hombre que sólo sea matemático y un artesano común. También la llamo la ocupación más bella del mundo; pero es sólo una ocupación; he dicho muchas veces que es bueno intentarlo [estudiar matemáticas], pero sin agotar nuestras fuerzas; así que yo no daría dos pasos por las matemáticas, y estoy seguro de que usted apoya firmemente mi opinión.» Pascal era hombre de cualidades muy variadas, aunque contradictorias.

Desargues urgió a Pascal a trabajar en el método de proyección y sección, sugiriéndole en particular el objetivo de reducir las muchas propiedades de la secciones cónicas a un pequeño número de proposiciones básicas. Pascal adoptó estas recomendaciones, y en 1639, a los dieciséis años, escribió un libro sobre cónicas que empleaba métodos proyectivos, es decir, proyección y sección. Este trabajo se ha perdido, pero Leibniz lo vio en París en 1676, y se lo describió al sobrino de Pascal. Un *Ensayo sobre las cónicas* (1640) de ocho páginas, conocido de unos pocos contemporáneos de Pascal, se perdió también, hasta 1779, en que fue recuperado [3]. Descartes, que leyó su ensayo de 1640, lo consideró tan brillante que no podía creer que lo hubiese escrito alguien tan joven.

El resultado más famoso de Pascal en geometría proyectiva, que aparece en las dos obras antes citadas, es un teorema que lleva hoy su nombre, y que en lenguaje moderno afirma lo siguiente: si se tiene un hexágono inscrito en una cónica, los tres puntos de intersección de los pares de lados opuestos están en línea recta. Así (fig. 14.12) P, Q y R están alineados. Si los lados opuestos del hexágono son paralelos, P, Q y R están en la recta del infinito.

Sólo tenemos indicaciones acerca de cómo probó Pascal este teorema. Afirma que, al ser cierto para el círculo, por proyección y

---

[3] *Oeuvres*, 1, 1908, 243-60.

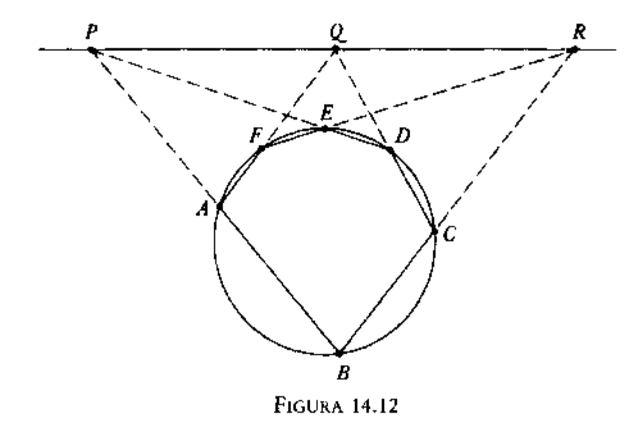


FIGURA 14.12

sección debe ser cierto para todas las cónicas. Y es claro que si uno efectúa una proyección de dicha figura desde un punto exterior al plano, seguida de una sección, ésta contendrá una cónica y un hexágono inscrito en ella. Además, los lados opuestos del hexágono se cortarán en tres puntos que se hallarán sobre una recta, puntos y recta homólogos a *P*, *Q*, *R* y la recta *PQR* de la figura original. Por cierto, el teorema de Pappus (cap. 5, sec. 7), que se refiere a tres puntos sobre dos rectas, es un caso particular de este teorema. Cuando la cónica degenera en dos rectas, como cuando una hipérbola degenera en sus asíntotas, se obtiene la situación descrita por Pappus.

El recíproco del teorema de Pascal, a saber: «si un hexágono es tal que los puntos de intersección de sus tres pares de lados opuestos están en línea recta, entonces sus vértices están sobre una cónica», es también cierto, aunque Pascal no lo consideró. Hay otros resultados en la obra de Pascal de 1640, pero no vale la pena analizarlos aquí.

El método de proyección y sección fue continuado por Philippe de La Hire (1640-1718), que fue pintor en su juventud y después se dedicó a las matemáticas y la astronomía. Como Pascal, La Hire fue muy influido por Desargues y llevó a cabo considerable trabajo sobre las secciones cónicas. Parte de él, aparecido en publicaciones de 1673 y 1679, empleaba la manera sintética de los griegos, pero con nuevos métodos, como las definiciones focales de la elipse y la hipérbola, y en algunas cosas utilizó la geometría analítica de Descartes y Fermat. Su trabajo más importante, sin embargo, es *Sectiones Conicae* (1685), y está dedicado a la geometría proyectiva. Como Desargues y Pascal, La Hire demostraba primero propiedades del círculo, relativas sobre

todo a cuaternas armónicas, y las trasladaba a otras secciones cónicas por proyección y sección. Podía así trasladar las propiedades del círculo a cualquier tipo de sección cónica con un solo método de demostración. Aunque hay algunas omisiones, como el teorema de involución de Desargues y el teorema de Pascal, hallamos en la obra de La Hire de 1685 prácticamente todas las propiedades de las cónicas que hoy nos son familiares, demostradas sintéticamente y establecidas sistemáticamente. De hecho, La Hire demuestra casi todos los 364 teoremas de Apolonio sobre las cónicas. También prueba las propiedades armónicas de los cuadriláteros. En total, La Hire demostró unos 300 teoremas. Trató de probar que los métodos proyectivos eran superiores a los de Apolonio y a los nuevos métodos analíticos de Descartes y Fermat (cap. 15), que ya habían sido creados.

Globalmente considerados, los resultados de La Hire no van más allá de los de Desargues y Pascal. Sin embargo, en la teoría de polos y polares obtiene un resultado nuevo importante: demuestra que si un punto describe una línea recta, entonces la polar del punto gira alrededor del polo de la recta. Así, si Q (fig. 14.13) se mueve a lo largo de la recta p, la polar de Q gira alrededor del polo P de la recta p.

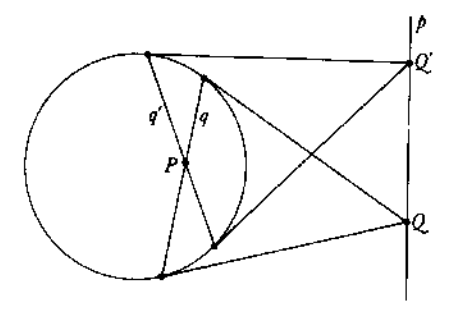


FIGURA 14.13

5. La aparición de nuevos principios

Por encima de los teoremas concretos de Desargues, Pascal y La Hire surgían nuevas ideas y perspectivas. La primera es la idea de *cambio continuo de una entidad matemática* de un estado a otro; en nuestro caso, la entidad es una figura geométrica. Fue Kepler, en su

Astronomia pars Optica[4] de 1604, el primero que pareció darse cuenta de que la parábola, la elipse, la hipérbola, el círculo y la cónica degenerada formada por un par de rectas se pueden derivar continuamente unas de otras. Para llevar a cabo este proceso de cambio continuo de una figura en otra, por ejemplo, de elipse a parábola y después a hipérbola, Kepler fija un foco y supone el otro móvil a lo largo de la recta que los une. Al hacer que el foco móvil se desplace hasta el infinito (y que la excentricidad se aproxime a 1), la elipse se convierte en parábola, y si ahora hacemos que el foco móvil reaparezca en la recta, pero al otro lado del foco estacionario, la parábola se convierte en hipérbola. Cuando los dos focos de una elipse se juntan, ésta se convierte en un círculo, y cuando los dos focos de una hipérbola se juntan, ésta degenera en dos rectas. Para hacer posible que un foco pueda alejarse hasta el infinito y luego reaparecer al otro lado, Kepler suponía, como ya hemos indicado, que la línea recta, extendida en ambas direcciones, se corta a sí misma en el infinito en un punto, de forma que la recta tiene las propiedades de un círculo. Aunque no es satisfactorio desde el punto de vista intuitivo visualizar así la recta, la idea es lógicamente sólida, y llegó a ser fundamental en la geometría proyectiva del siglo XIX. Kepler indicó también que las diversas secciones del cono pueden obtenerse variando continuamente la inclinación del plano que las produce.

La noción de cambio continuo en una figura es también empleada por Pascal. Haciendo converger dos vértices consecutivos de un hexágono, éste se transforma en pentágono, lo cual le permite razonar sobre propiedades del pentágono a partir de las correspondientes del hexágono, analizando lo que les sucede ante cambios continuos. Análogamente se pasa de pentágonos a cuadriláteros.

La segunda idea que surge claramente de la obra de los geómetras «proyectivos» es la de *transformación e invariancia*. Proyectar una figura desde un cierto punto y luego obtener una sección de dicha proyección es transformar la figura original en otra nueva. Las propiedades de ésta que nos interesan son aquellas que permanecen invariantes por la transformación. Otros geómetras del siglo XIX, como Gregorio de St. Vincent (1584-1667) y Newton[5], introdujeron transformaciones distintas de la proyección y sección.

[4] *Ad Vitellionem Paralipomena, quibus Astronomiae pars Optica Traditur* (Suplemento a Vitello, incluyendo la parte óptica de la astronomía).

[5] *Principia*, 3. ed., libro 1, lema 22 y prop. 25.

Estos geómetras proyectivos continuaron además la labor de búsqueda de métodos generales iniciada por los algebristas, Vieta en particular. En la época griega, los métodos de demostración habían sido de potencia limitada. Cada teorema exigía un plan de ataque distinto. Ni Euclides ni Apolonio parecen haberse preocupado de los métodos generales. Sin embargo, Desargues insistió en el método de proyección y sección porque veía en él un procedimiento general para probar teoremas sobre cónicas cualesquiera, una vez demostrados para el círculo. Además, vio en las nociones de involución y cuaterna armónica conceptos más generales que los de la geometría euclídea. En efecto, una cuaterna armónica, cuando uno de sus puntos está en el infinito, se reduce a tres puntos, uno de los cuales es el punto medio de los otros dos, por lo cual la noción de cuaterna armónica y los teoremas relativos a ellas son más generales que la noción de punto bisector de un segmento. Desargues y Pascal trataban de deducir el máximo número posible de resultados de un solo teorema. Bosse indica que Desargues dedujo sesenta teoremas de Apolonio a partir de su teorema de involución, y que Pascal le felicitó por ello. Este, en su búsqueda de las relaciones entre figuras distintas, como hexágono y pentágono, trató de hallar una aproximación común a ambas. Se piensa, en efecto, que dedujo unos 400 corolarios de su teorema del hexágono a base de analizar sus consecuencias para figuras relacionadas con él, aunque no se ha hallado ninguna obra suya en la que esto aparezca. El interés por el método es evidente en la obra de La Hire de 1685, cuyo objetivo principal era precisamente probar que el método de proyección y sección era superior a los de Apolonio, e incluso a los métodos algebraicos de Descartes. Este afán por la generalidad en resultados y métodos se convirtió en fuerza poderosa en la matemática posterior.

Mientras los geómetras insistían en la generalidad del método, estaban descubriendo inconscientemente otro tipo de generalidad. Muchos de los teoremas, como el del triángulo de Desargues, tratan de intersecciones de rectas, y no de longitudes, ángulos y áreas, como sucede en la geometría euclídea. La intersección de dos rectas es lógicamente anterior a cualquier consideración de tamaño, pues el propio hecho de la intersección determina la formación de una figura. Estaba naciendo una nueva y fundamental rama de la geometría interesada en las propiedades de posición e intersección, y no en las métricas o de tamaño. Sin embargo, los matemáticos del siglo XVII que trabajaron en geometría proyectiva empleaban la geometría eucli-

dea como base, particularmente los conceptos de distancia y medida de ángulos. Además, estos geómetras, lejos de pensar en términos de la nueva geometría, intentaban de hecho mejorar los métodos de la geometría euclídea. La conciencia de que en su trabajo se hallaba implícita una nueva rama de la geometría no llegó hasta el siglo XIX.

Si bien la motivación para el estudio de la geometría proyectiva fue originalmente el deseo de ayudar a los pintores, este objetivo se desvió y se fundió con el creciente interés por las secciones cónicas. Pero la matemática pura no resultaba agradable al espíritu del siglo XVII, cuyos matemáticos estaban mucho más interesados en entender y dominar la naturaleza; es decir, en los problemas científicos. Los métodos algebraicos resultaron ser más efectivos en el tratamiento de los problemas matemáticos, y proporcionaban el conocimiento cuantitativo que ciencia y tecnología buscaban. Los resultados cualitativos obtenidos por la geometría proyectiva con sus métodos sintéticos no eran, ni mucho menos, tan útiles. De aquí que fuese abandonada en favor del álgebra, la geometría analítica y el cálculo, que dieron lugar, independientemente, a nuevas áreas de central importancia en la matemática moderna. Los resultados de Desargues, Pascal y La Hire fueron olvidados y hubieron de redescubrirse más tarde, en el siglo XIX sobre todo, época en la cual las creaciones y nuevos puntos de vista surgidos entre tanto permitieron a los matemáticos sacar el fruto de las grandes ideas aún latentes en la geometría proyectiva.

Bibliografía

Chasles, Michel: *Aperçu historique des méthodes en géometrie,* 3 ed., Gauthier-Villars et Fils, pp. 68-95, 118-37 (igual que en la primera edición de 1837).

Coolidge, Julian L.: *A History of Geometrical Methods.* Dover (reimpresión), 1963, pp. 88-92.

—: *A History of the Conic Sections and Quadric Surfaces,* Dover (reimpresión), 1968, cap. 3.

—: «The Rise and Fall of Projective Geometry», *Amer. Math. Monthly,* 41, 1934, 217-28.

Desargues, Girard: *Oeuvres,* 2 vols., Leiber, 1864.

Ivins, W. M., Jr.: «A Note on Girard Desargues», *Scripta Math.,* 9, 1943, 33-48.

—: «A Note on Desargues's Theorem», *Scripta Math.,* 13, 1947, 203-10.

Mortimer, Ernest: *Blaise Pascal: The Life and Work of a Realist*, Harper and Bros., 1959.
Pascal, Blaise: *Oeuvres,* Hachette, 1914-21.
—: *Obras. 1,* Madrid, Alfaguara,l 1981.
Smith, David Eugene: *A Source Book in Mathematics,* Dover (reimpresión), 1959, vol. 2, pp. 307-14, 326-30.
Struik, D. J.: *A Source Book in Mathematics, 1200-1800,* Harvard University Press, 1969, pp. 157-68.
Taton, René: *L'Oeuvre mathématique de G. Desargues,* Presses Universitaires de France, 1951.

Capítulo 15
LA GEOMETRIA ANALITICA

> He decidido abandonar la geometría abstracta, es decir, la consideración de cuestiones que *sólo sirven para ejercitar la mente*, para estudiar otro tipo de geometría que tiene por objeto la explicación de los fenómenos de la naturaleza.
>
> René Descartes

1. La motivación de la geometría de coordenadas

Fermat y Descartes, los dos principales responsables de la gran creación matemática que siguió a las que hemos estudiado, estaban interesados, como Desargues y sus seguidores, en hallar métodos generales para el estudio de curvas. Pero Fermat y Descartes se hallaban inmersos en el trabajo científico, siendo, por ello, plenamente conscientes de la necesidad de métodos cuantitativos, y la impresión que causó en ambos la potencia del álgebra hizo que se volcaran a la aplicación de ésta al estudio de la geometría. La disciplina que crearon se llama geometría de coordenadas o analítica, y su idea central es asociar ecuaciones algebraicas a las curvas y superficies. Es ésta una de las vetas más ricas y fructíferas del pensamiento matemático que jamás se hayan encontrado.

Está fuera de toda duda que la motivación de Fermat y Descartes se hallaba en las necesidades de la ciencia y en su interés por la metodología. Las aportaciones de Fermat al cálculo, la construcción de tangentes y la obtención de máximos y mínimos fueron motivadas, como veremos más claramente al hablar de la historia del cálculo, por la resolución de problemas científicos. Hizo, además, aportaciones de

primer orden a la óptica. Su interés por la metodología puede comprobarse explícitamente en su breve libro *Ad Locos Planos et Solidos Isagoge* (Introducción a los lugares planos y sólidos [1]), escrito en 1629, pero publicado en 1679 [2], en el que habla de su búsqueda de métodos universales para los problemas de curvas. Fue, al igual que Descartes, uno de los científicos más grandes de siglo XVII, y como él, hizo del método un objetivo primario de todo su trabajo.

2. La geometría analítica de Fermat

En sus trabajos en teoría de números, Fermat partió de la obra de Diofanto. Para sus investigaciones sobre curvas, partió de la obra de los geómetras griegos, sobre todo Apolonio, cuyo libro perdido *Lugares planos* había reconstruido. Gracias a sus contribuciones al álgebra, estaba en condiciones de aplicarla al estudio de las curvas, cosa que hizo en *Ad Locos*, afirmando su propósito de inaugurar un estudio general de los lugares geométricos que los griegos no habían llegado a hacer. No conocemos la manera exacta en que evolucionaron las ideas de Fermat sobre la geometría analítica. Aunque estaba familiarizado con el uso que daba Vieta al álgebra en la resolución de problemas geométricos, lo más probable es que tradujese directamente los resultados de Apolonio a una forma algebraica.

Fermat considera una curva cualquiera y un punto genérico J sobre ella (fig. 15.1). La posición de J viene fijada por una longitud A, medida desde un punto O sobre una línea de base a un punto Z, y la longitud E de Z a J. Fermat emplea así lo que hoy llamamos coordenadas oblicuas, aunque no aparece explícitamente ningún eje de las y, ni se emplean coordenadas negativas. Sus A y E son nuestras x e y.

Fermat había enunciado con anterioridad su principio general: «Siempre que en una ecuación se hallen dos cantidades incógnitas, tenemos un lugar geométrico, cuyo extremo describe una línea recta o curva.» Así, los extremos J, J', J'', ... de E en sus diversas posiciones describen la «línea». Sus cantidades desconocidas A y E son en realidad variables, esto es, la ecuación en A y E es indeterminada. Aquí usa Fermat la idea de Vieta de representar con una letra toda una

[1] Fermat utiliza estos términos en el sentido de Pappus. Ver cap. 8, sec. 2.
[2] *Oeuvres*, 1, 91-103.

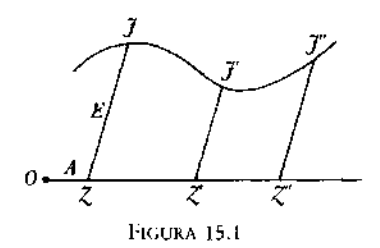


FIGURA 15.1

clase de números. A continuación, Fermat da varias ecuaciones algebraicas en *A* y *E* y analiza las curvas que describen. Así, escribe «*D in A aequetur B in E*» (en nuestra notación, $Dx = By$), y afirma que representa una línea recta. Da también la ecuación más general $d(a - x) = by$ en nuestra notación, concluyendo que también representa una recta. La ecuación «*B quad. - A quad. aequetur E quad*» (en nuestra notación, $B^2 - x^2 = y^2$) representa un círculo. Análogamente, $a^2 - x^2 = ky^2$ representa una elipse, $a^2 + x^2 = ky^2$ y $xy = a$ representan hipérbolas, y $x^2 = ay$ representa una parábola. Como Fermat no usaba coordenadas negativas, sus ecuaciones no podían representar las curvas completas que pretendía describir. Era consciente de la posibilidad de trasladar y girar los ejes, ya que consideró ecuaciones más complicadas de segundo grado, dando las formas más simples a las que se pueden reducir. Afirma, de hecho, que una ecuación de primer grado en *A* y *E* tiene como lugar geométrico una recta, y todas las ecuaciones de segundo grado tienen cónicas como lugares geométricos. En su *Methodus ad Disquiriendam Maximam et Minimam* (Método para hallar máximos y mínimos, 1637) [3], introdujo las curvas de $y = x^n$ e $y = x^{-n}$.

3. René Descartes

Descartes fue el primer gran filósofo moderno, uno de los fundadores de la biología moderna, un físico de primera fila y, sólo incidentalmente, matemático. Pero cuando un hombre de tal potencia intelectual dedica aunque sólo sea una parte de su tiempo a un tema, los resultados han de ser forzosamente importantes.

---

[3] *Oeuvres*, 1, 133-79; 3, 121-56.

Nació en La Haye, Turena (Francia), el 31 de marzo de 1596. Su padre, un abogado moderadamente rico, le envió a los ocho años de edad al colegio de jesuitas de La Flèche, en Anjou. Por su delicada salud, le era permitido pasar las mañanas en la cama, tiempo que aprovechaba para estudiar, y conservó esa costumbre durante toda su vida. A los dieciséis años dejó La Flèche y a los veinte se licenció en la Universidad de Poitiers con el título de abogado y marchó a París. Allí encontró a Mydorge y al P. Marin Mersenne y pasó un año con ellos estudiando matemáticas. Descartes, sin embargo, era muy inquieto y se enroló en el ejército del príncipe Mauricio de Orange en 1617. Pasó los nueve años siguientes alternando el servicio en diversos ejércitos con la juerga en París, aunque siguió estudiando matemáticas en todo este período. Su habilidad para resolver un problema anunciado en un tablón en Breda (Holanda) le convenció de su capacidad matemática, y empezó a plantearse seriamente dedicarse a ello. Volvió a París, y, entusiasmado por la potencia del telescopio, se recluyó para estudiar la teoría y construcción de instrumentos ópticos. En 1628 se fue a Holanda en busca de ambiente intelectual más libre y seguro. Allí vivió veinte años y escribió sus famosas obras. En 1649 fue invitado a la corte de la reina Cristina de Suecia. Aceptó, tentado por el honor y el atractivo de la realeza, y allí murió de neumonía en 1650.

Su primera obra, *Regulae ad Directionem Ingenii* (Reglas para la dirección del espíritu) [4], fue escrita en 1628 y publicada póstumamente. Su siguiente obra importante fue *Le Monde* (Sistema del mundo, 1634), que contiene una teoría cosmológica de vórtices para explicar cómo se mantienen los planetas en su movimiento propio y en sus órbitas alrededor del sol. Sin embargo, no la publicó por miedo a ser perseguido por la Iglesia. En 1637 publicó el *Discours de la méthode pour bien conduire la raison, et chercher la vérité dans les sciences* [5]. Este libro, un clásico de la literatura y la filosofía, contiene tres famosos apéndices: *La Géometrie, La Dioptrique* y *Les Méteores. La Géometrie* contiene sus ideas sobre la geometría analítica y el álgebra, y es el único libro de matemáticas escrito por Descartes, aunque no dejó de comunicar por carta otras muchas ideas matemáticas. El *Discours* le trajo gran fama de inmediato, y su entusiasmo y el de su público no dejaron de crecer con el paso del tiempo. En 1644 publicó *Principia Philosophiae*, dedicada a la ciencia física y en especial a las

[4] Publicado en holandés en 1692; *Oeuvres*, 10, 359-469.

[5] *Oeuvres*, 6, 1-78.

leyes del movimiento y la teoría de vórtices. Contiene material de su *Sistema*, que ahora consideraba más aceptable para la Iglesia. En 1650 publicó *Musicae Compendium*.

Las ideas de Descartes dominaron el siglo XVII. Sus enseñanzas y escritos llegaron a ser conocidos incluso entre personas ajenas a la ciencia, por la forma tan atractiva en que él las presentaba. Sólo la Iglesia le rechazó. En realidad, Descartes era persona devota, y feliz de haber establecido (como creía) la existencia de Dios. Pero había enseñado que la Biblia no era la fuente del conocimiento científico, que la razón por sí sola bastaba para establecer la existencia de Dios, y que el hombre sólo debía aceptar aquello que pudiese entender. La Iglesia reaccionó ante tales enseñanzas incluyendo el libro en el *Indice de libros prohibidos* poco después de su muerte, y prohibiendo las oraciones fúnebres con ocasión de su entierro en París.

Descartes llegó a la matemática por tres vías: la filosofía, el estudio de la naturaleza y el interés por los usos de la ciencia. Es difícil, y quizá artificial, tratar de separar estas tres líneas de pensamiento. Vivió la culminación de la controversia protestante y los primeros logros de la ciencia en el descubrimiento de leyes de la naturaleza que desafiaban a muchas doctrinas religiosas. Todo ello hizo que Descartes empezase a dudar de los conocimientos que había adquirido en la escuela. Al término de sus estudios en La Flèche, había llegado ya a la conclusión de que su educación no había hecho más que contribuir a su perplejidad. Se vio asediado por dudas, y se convenció de que no había más progreso que el reconocimiento de su propia ignorancia. Pero como había asistido a una de las escuelas más famosas de Europa, y no se consideraba mal estudiante, se vio justificado para dudar de que hubiera en alguna parte algún cuerpo de conocimiento seguro. Se planteó así la cuestión: ¿Cómo podemos llegar al conocimiento de algo?

Pronto llegó a la conclusión de que la lógica en sí misma era estéril: «En cuanto a la Lógica, los silogismos y los demás preceptos son de utilidad para la comunicación de lo que ya sabemos, o... incluso para hablar sin juicio de cosas que ignoramos, pero no para investigar lo desconocido.» La lógica, pues, no nos proporciona las verdades fundamentales.

¿Y dónde hallarlas? Rechazó la filosofía del momento, fundamentalmente escolástica, que, aunque sugestiva, parecía no disponer de fundamentos claros y cuyos razonamientos no siempre eran irreprochables. La filosofía, concluyó, sólo proporciona «los medios para

disertar sobre cualquier materia con apariencia de verdad». La teología apuntaba hacia un cielo al que aspiraba a llegar tanto como cualquier otro, pero ¿era correcto el camino?

El método para establecer la verdad en cualquier campo se le ocurrió, según él, en un sueño, el 10 de noviembre de 1619, durante una de sus campañas militares. Era el método de las matemáticas. Estas le atraían porque las pruebas basadas en sus axiomas eran irrefutables, y la autoridad no contaba para nada. Las matemáticas proporcionaban el método para llegar a las verdades y para demostrarlas efectivamente, y era evidente que dicho método trascendía a su materia de estudio. Dice: «Es un método de conocimiento más potente que ningún otro que nos haya sido otorgado por obra humana, y es la fuente de todos los demás.» Continúa en la misma idea:

> ...Todas las ciencias que tienen como fin la investigación sobre el orden y la medida están relacionadas con las matemáticas, y poco importa si esa medida se busca en los números, formas, estrellas, sonidos o cualquier otro objeto; por todo ello, debe existir una ciencia general que explique todo lo que deba ser conocido sobre el orden y la medida, con independencia de su aplicación a alguna disciplina particular, y es así que esta ciencia tiene su propio nombre, consagrado por su prolongado uso, y es el de matemáticas. Y una prueba de que sobrepasa con mucho en facilidad e importancia a las ciencias que de ella dependen es que abarca a la vez todos los objetos a las que éstas se dedican, además de muchos otros...

Y así concluye que «las largas cadenas de razonamientos simples y fáciles a que están acostumbrados los geómetras para alcanzar las conclusiones de sus más difíciles demostraciones me han llevado a imaginar que todas las cosas cuyo conocimiento compete al hombre están mutuamente relacionadas de la misma forma».

De su estudio del método matemático extrajo en las *Reglas para la dirección del espíritu* los siguientes principios para asegurar la exactitud del conocimiento en cualquier campo: no aceptar como verdadero nada que no esté en la mente de forma tan clara y distinta que excluya cualquier duda; dividir las dificultades en otras menores; proceder de lo simple a lo complejo; y, finalmente, enumerar y revisar los pasos del razonamiento de forma tan completa que nada pueda omitirse.

Con estos elementos esenciales del método, destilados de su práctica matemática, Descartes esperaba resolver cuestiones de filosofía, física, anatomía, astronomía, matemáticas y otros campos. Y

aunque no tuvo éxito en tan ambicioso programa, sí hizo contribuciones notables a la filosofía, las ciencias y las matemáticas. La aprehensión inmediata por la mente de verdades básicas, claras y distintas, el poder de la intuición y la deducción de consecuencias son la esencia de su filosofía del conocimiento. Cualquier pretensión de conocimiento obtenido de otra forma debe rechazarse como sospechosa de error y peligrosa. El propósito de los tres apéndices del *Discours* es mostrar la efectividad de su método, cosa que creyó haber logrado.

Descartes inauguró la filosofía moderna. No podemos aquí seguir hablando de su sistema, sino sólo de unos pocos aspectos de importancia para las matemáticas. En filosofía, tomó como axiomas verdades que le pareciesen tan claras como para aceptarlas de inmediato, adoptando finalmente cuatro: (a) *cogito, ergo sum* (pienso, luego existo); (b) todo fenómeno tiene una causa; (c) un efecto no puede ser mayor que su causa; (d) son innatas a la mente las ideas de perfección, espacio, tiempo y movimiento. La idea de perfección, de ser perfecto, no puede ser deducida o creada por la mente imperfecta del hombre, en virtud del axioma (c). Sólo puede obtenerse a partir de un ser perfecto. Luego Dios existe. Como Dios no puede engañarnos, podemos estar seguros de que los axiomas de las matemáticas, que son claros para nuestra intuición, y las deducciones que de ellos hacemos mediante procesos puramente mentales se aplican realmente al mundo físico y son, pues, verdades. Se sigue así que Dios tiene que haber establecido la naturaleza según leyes matemáticas.

En cuanto a las propias matemáticas, creía en las ideas claras y distintas, como la de triángulo. Estas ideas matemáticas existen y son eternas e inmutables, y no dependen de que se piense en ellas o no. Las matemáticas, pues, tienen una existencia externa y objetiva.

La segunda cuestión esencial para Descartes, y para casi todos los pensadores de su época, era entender la naturaleza. Dedicó muchos años a los problemas científicos y experimentó intensamente en mecánica, hidrostática, óptica y biología. Su teoría de los vórtices fue la dominante en cosmología durante todo el siglo XVII. Fue el fundador de la filosofía mecanicista, según la cual todos los fenómenos naturales, incluyendo el funcionamiento del cuerpo humano y excluyendo el alma, se reducen a movimientos que obedecen a las leyes de la mecánica. La óptica, particularmente el diseño de lentes, tenía gran interés para él. Parte de *La Géometrie* está dedicada a la óptica, y lo mismo sucede con *La Dioptrique*. Descartes comparte con Willebrord Snell el honor del descubrimiento de la ley de

refracción de la luz. Como en filosofía, su obra científica fue fundamental y revolucionaria.

De gran importancia en esta obra científica es el interés de Descartes para poner en uso los frutos de la ciencia (cap. 11, sec. 5), en lo cual se separa clara y abiertamente de los griegos. Dominar la naturaleza para el bien del hombre es la motivación de muchos de los problemas científicos a los que se dedicó. E impresionado por el poder de las matemáticas, se aplicó en utilizarlas; para él, no era una disciplina contemplativa, sino una ciencia constructiva y útil. Al contrario que a Fermat, poco le importaba su belleza y armonía, y menos aún valoraba la matemática pura, llegando a afirmar que el método matemático aplicado sólo a las matemáticas carece de valor por no ser parte del estudio de la naturaleza. Aquellos que cultivan las matemáticas por sí mismas son investigadores ociosos entregados a un juego vano del espíritu.

4. La obra de Descartes en geometría analítica

Convencido de la importancia del método y de la posibilidad de emplear con provecho las matemáticas en el trabajo científico, Descartes se dedicó a aplicar dicho método a la geometría. En esa empresa juntaron sus fuerzas su interés general en el método y su particular conocimiento del álgebra. Le molestaba que cada demostración de la geometría euclídea exigiese argumentos nuevos y a menudo ingeniosos. Tildaba explícitamente a la geometría de los clásicos de abstracta en exceso y tan ligada a las figuras «que sólo puede ejercitar el entendimiento a condición de fatigar grandemente la imaginación». El álgebra, que tan extendida halló, fue también objeto de su crítica, por estar tan completamente sujeta a reglas y fórmulas «que resulta un arte lleno de confusión y oscuridad e idóneo para estorbar, y no una ciencia útil para el desarrollo de la mente». Por todo ello, Descartes propone tomar lo mejor del álgebra y la geometría y corregir los defectos de una con la ayuda de la otra.

Lo que en realidad emprendió fue la aplicación del álgebra a la geometría. Vio con claridad la potencia del álgebra y su superioridad sobre los métodos geométricos de los griegos en la creación de una metodología más amplia. Hizo asimismo notar la generalidad del álgebra y su valor en la mecanización del proceso del razonamiento y en la reducción del trabajo en la resolución de problemas, y percibió

su potencial como ciencia universal del método. El producto de esta aplicación del álgebra a la geometría fue *La Géometrie*.

Aunque Descartes usó las mejoras en la notación algebraica que ya hemos indicado en el capítulo 13, su libro no es fácil de leer. Esta oscuridad era en gran parte deliberada, y Descartes presumía de que pocos matemáticos en Europa podían entender su trabajo. Se limitó a indicar las construcciones y demostraciones, dejando a otros la labor de completar los detalles. En una carta compara su manera de escribir con la de un arquitecto que traza los planos y prescribe lo que ha de hacerse, y deja el trabajo manual a carpinteros y albañiles. Dice también: «Nada he omitido por descuido, más preveo que ciertas personas que presumen de saberlo todo no perderían la oportunidad de decir que no he escrito nada que ellos no conociesen si me expresase de forma suficientemente inteligible para ellos.» Muchos comentarios explicativos se han escrito para hacer comprensible el libro de Descartes.

Sus ideas sólo pueden entenderse mediante los ejemplos que da en el libro. Omite la demostración de la mayor parte de sus enunciados generales alegando que si uno se molesta en examinar sistemáticamente estos ejemplos, las demostraciones de los resultados generales resultan claras, y es mejor aprenderlas de esta manera.

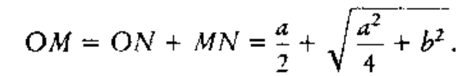
Comienza *La Géometrie* con la resolución de problemas geométricos mediante el álgebra, a la manera de Vieta, y sólo de forma gradual va surgiendo la idea de la ecuación de una curva. Apunta en primer lugar que las construcciones geométricas requieren sumar, restar, multiplicar y dividir líneas, así como extraer raíces cuadradas de líneas concretas. Como todas estas operaciones existen también en el álgebra, pueden expresarse en términos algebraicos.

Para abordar un problema dado, Descartes dice que debemos suponer conocida la solución y representar con letras todas las líneas, conocidas y desconocidas, que sean necesarias para la construcción buscada. Así, al no hacer distinción entre líneas conocidas y desconocidas, hemos de «desenmarañar» la dificultad mostrando en qué forma están relacionadas entre sí estas líneas, con la idea de expresar la misma cantidad de dos formas distintas, lo cual nos da una ecuación. Hemos de encontrar tantas ecuaciones como líneas desconocidas haya. Si quedan varias ecuaciones, hay que combinarlas hasta que sólo quede una línea incógnita expresada en términos de líneas conocidas. A continuación, Descartes muestra cómo construir la línea desconocida empleando el hecho de que satisface la ecuación algebraica.

Supongamos, por ejemplo, que un problema geométrico nos lleva a hallar una longitud desconocida x, y que la formulación algebraica indica que x debe satisfacer la ecuación $x^2 = ax + b^2$, donde a y b son longitudes conocidas. Sabemos entonces por álgebra que

$$x = \frac{a}{2} + \sqrt{\frac{a^2}{4} + b^2} \qquad (1)$$

(Descartes no tenía en cuenta la segunda raíz, que es negativa.) Ahora Descartes dice cómo construir x. Construye un triángulo rectángulo NLM (fig. 15.2) con $LM = b$ y $NL = a/2$, y prolonga MN hasta O de forma que $NO = NL = a/2$. Entonces, la solución x es la longitud OM. Descartes no da la prueba de que OM es la longitud correcta, que es inmediata, pues



$$OM = ON + MN = \frac{a}{2} + \sqrt{\frac{a^2}{4} + b^2}.$$

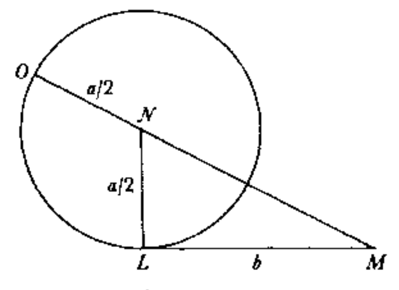


FIGURA 15.2

Por tanto, la expresión (1) de x, obtenida resolviendo una ecuación algebraica, indica la construcción apropiada de x.

En la primera mitad del libro I, Descartes resuelve exclusivamente problemas clásicos con ayuda del álgebra, lo que constituye una aplicación del álgebra a la geometría, pero no geometría analítica en nuestro sentido actual. Hasta este punto sólo ha considerado problemas de construcción determinados, que tienen una única solución. Aborda a continuación problemas indeterminados, en los que hay muchas longitudes como soluciones posibles. Los extremos de todas estas longitudes rellenan una curva, y aquí, según Descartes, «se pide también descubrir y trazar la curva que contiene todos esos puntos».

Esta curva viene descrita por la ecuación indeterminada final que expresa las longitudes incógnitas y en términos de las longitudes arbitrarias x. Descartes insiste además en que para cada x, y satisfaga una ecuación determinada, de forma que pueda construirse. Si la ecuación es de primer o segundo grado, y puede construirse por los métodos del libro I, por medio de rectas y círculos exclusivamente. Para ecuaciones de grado superior, anuncia que mostrará su construcción en el libro III.

Descartes usa el teorema de Pappus (cap. 5, sec. 7) para ilustrar lo que sucede cuando un problema da lugar a una ecuación con dos incógnitas. Este problema, que no había sido resuelto en toda su generalidad, dice lo siguiente: dadas tres rectas en un plano, hallar la posición (o lugar geométrico) de todos los puntos desde los que se pueden trazar tres rectas, una para cada una de las dadas, que formen un ángulo determinado con cada una de ellas (el ángulo puede variar de una recta a otra), de forma que el rectángulo formado por dos de las líneas construidas esté en una razón dada con el cuadrado de la tercera línea construida; si son cuatro las rectas dadas inicialmente, entonces las rectas construidas formando ángulos dados con las dadas deben ser tales que el rectángulo formado por dos de ellas debe estar en una razón dada con el formado por las otras dos; si hay inicialmente cinco rectas, entonces las cinco rectas construidas formando un ángulo dado con cada una de las anteriores deben ser tales que el producto de tres de ellas esté en una razón dada con el producto de las dos restantes. La condición exigida al lugar geométrico cuando se dan más de cinco líneas es una extensión obvia de la ya mencionada.

Pappus había anunciado que cuando se dan tres o cuatro rectas, el lugar buscado es una sección cónica. En el libro II, Descartes trata el problema de Pappus para el caso de cuatro rectas. Estas son (fig. 15.3)

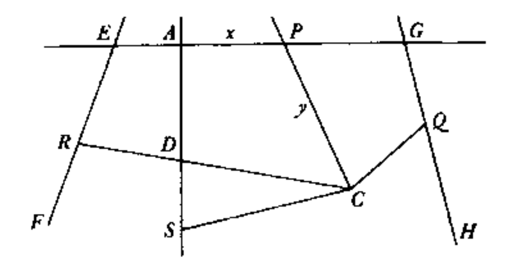


FIGURA 15.3

AG, GH, EF y AD. Consideremos un punto C y las cuatro rectas desde C a cada una de las cuatro rectas dadas y que forman con cada una de ellas un ángulo prefijado, que puede variar de una recta a otra. Denotemos estas cuatro rectas por CP, CQ, CR y CS. Se pide hallar el lugar de los puntos C que satisfacen la condición $CP \cdot CR = CS \cdot CQ$.

Descartes llama x a AP e y a PC. Por consideraciones geométricas sencillas, obtiene los valores de CR, CQ y CS en términos de cantidades conocidas. Emplea estos valores en la condición $CP \cdot CR = CS \cdot CQ$ y obtiene una ecuación de segundo grado en x e y de la forma

$$y^2 = Ay + Bxy + Cx + Dx^2 \qquad (2)$$

donde A, B, C y D son expresiones algebraicas sencillas en términos de las cantidades conocidas. Ahora indica Descartes que si seleccionamos un valor cualquiera de x, tenemos una ecuación cuadrática en y de la que y puede despejarse, y, por tanto, construirse con regla y compás, como ya se ha mostrado en el libro I. Por consiguiente, si se toman infinitos valores de x se obtienen un número infinito de valores de y, y, por tanto, un número infinito de puntos C. El lugar de todos estos puntos C es una curva cuya ecuación es (2).

Lo que ha hecho Descartes es establecer una recta (AG en la figura anterior) como línea de base con un origen en el punto A. Los valores x son, pues, longitudes a lo largo de esta recta, y los valores y son longitudes que comienzan en la línea de base y forman un ángulo fijo con ella. Este sistema de coordenadas es lo que ahora llamamos un sistema oblicuo. Las x e y de Descartes sólo pueden tomar valores positivos; sin embargo, sus ecuaciones cubren porciones de la curva en zonas distintas de la que hoy llamaríamos primer cuadrante. Se limita a suponer que el lugar geométrico en consideración se halla en el primer cuadrante, haciendo referencias de pasada sobre lo que podría suceder en otro caso. Que existe una longitud para cada número real positivo se da por hecho inconscientemente.

Una vez que ha llegado a la idea de la ecuación de una curva, Descartes procede a desarrollarla. Se demuestra fácilmente, afirma, que el grado de una curva es independiente de la elección del eje de referencia, y aconseja elegir este eje de forma que la ecuación resultante sea lo más sencilla posible. En otro gran paso adelante, considera dos curvas distintas, expresa sus ecuaciones con respecto a los mismos

ejes de referencia y halla sus puntos de intersección resolviendo las ecuaciones simultáneamente.

También en el libro II, Descartes considera críticamente las distinciones que hacían los griegos entre curvas planas, sólidas y lineales. Estos habían llamado curvas planas a las constructibles con regla y compás; las curvas sólidas eran las secciones cónicas, y las curvas lineales eran todas las demás, como la concoide, la espiral, la cuadratriz y la cisoide. Los griegos también llamaban a las lineales curvas mecánicas, porque se necesita cierto mecanismo especial para construirlas. Pero, dice Descartes, incluso la recta y el círculo requieren algún instrumento. Y, además, no nos puede importar la precisión de la construcción mecánica, pues lo único que cuenta en matemáticas es el razonamiento. Posiblemente, prosigue, los antiguos ponían objeciones a estas curvas por estar definidas de forma insegura. Sobre estas bases rechaza Descartes la idea de que sólo son legítimas las curvas constructibles con regla y compás [6], y llega a proponer nuevas curvas engendradas por construcciones mecánicas. Concluye con la muy significativa afirmación de que las curvas geométricas son las que pueden expresarse mediante una única ecuación algebraica (de grado finito) en x e y, con lo que acepta la concoide y la cisoide, mientras que llama mecánicas a todas las demás curvas, como la espiral y la cuadratriz.

Esta insistencia de Descartes en que las curvas aceptables son las que admiten una ecuación algebraica es el principio de la eliminación de la constructibilidad como criterio de existencia. Leibniz fue más lejos que Descartes: empleando las palabras «algebraica» y «trascendente» en vez de los términos de Descartes «geométrica» y «mecánica», protestaba contra el requerimiento de que una curva hubiese de tener una ecuación algebraica [7]. De hecho, Descartes y sus contemporáneos hicieron caso omiso de tal restricción y trabajaron con igual entusiasmo con la cicloide, la curva logarítmica, la espiral logarítmica ($\log \varrho = a\theta$), y otras curvas no algebraicas.

Al ampliar el concepto de curva admisible, Descartes dio un paso fundamental. No sólo admitía curvas anteriormente rechazadas, sino que ensanchaba su dominio, pues dada cualquier ecuación algebraica en x e y, puede obtenerse su curva, y generar así curvas totalmente nuevas. En *Arithmetica Universalis,* Newton dice (1707): «Pero los

[6] Compárese con la discusión del cap. 8, sec. 2.

[7] *Acta Erud.*, 1684, pp. 470, 587; 1686, p. 292 = *Math. Schriften,* 5, 127, 223, 226.

modernos geómetras, avanzando todavía mucho más allá [de las curvas planas, sólidas y lineales de los griegos], han recibido en la geometría todas las líneas que pueden expresarse por ecuaciones.»

Descartes prosigue considerando las clases de curvas geométricas. La primera y más simple es la formada por las de primer y segundo grados en x e y. Descartes dice en este sentido que las ecuaciones de las secciones cónicas son de segundo grado, pero no lo demuestra. Las curvas cuyas ecuaciones son de tercer y cuarto grados constituyen la segunda clase, las de grados quinto y sexto, la tercera, y así sucesivamente. La razón para agrupar los grados tercero y cuarto, así como el quinto y el sexto, es que creía que se podía reducir una del grado superior al inferior, de la misma forma en que las ecuaciones de cuarto grado pueden resolverse mediante las de tercero. Esta creencia era, por supuesto, incorrecta.

El tercer libro de *La Géometrie* vuelve al tema del libro I. Su objetivo es la resolución de problemas de construcciones geométricas que, al formularse algebraicamente, dan lugar a ecuaciones determinadas de grado tercero o superior y que, de acuerdo con el álgebra, precisan de las secciones cónicas y curvas de grado superior. Considera, por ejemplo, el problema de construir las dos medias proporcionales entre dos cantidades dadas a y q. El caso especial $q = 2a$ fue abordado muchas veces por los griegos clásicos por su importancia en el problema de la duplicación del cubo. Descartes procede como sigue: sea z una de las medias proporcionales; entonces z^2/a tiene que ser la segunda, ya que se ha de cumplir que

$$\frac{a}{z} = \frac{z}{z^2/a} = \frac{z^2/a}{z^3/a^2}.$$

Si ahora hacemos z^3/a^2 igual a q, obtenemos la ecuación que ha de satisfacer z. Dados q y a, hemos de hallar z tal que

$$z^3 = a^2q, \tag{3}$$

es decir, hemos de resolver una ecuación cúbica. Descartes muestra entonces que z y z^2/a pueden obtenerse por una construcción geométrica que utiliza una parábola y un círculo.

Tal y como describe Descartes la construcción, no aparecen coordenadas. Sin embargo, la parábola no se puede construir con

regla y compás sino punto a punto, y se debe, por tanto, utilizar la ecuación para dibujar la curva exactamente.

Descartes *no* obtiene z por el procedimiento de escribir las ecuaciones en x e y del círculo y la parábola y hallar las coordenadas del punto de intersección resolviendo el sistema de ecuaciones. En otras palabras, no resuelve las ecuaciones gráficamente en el sentido que le damos nosotros. En su lugar, usa construcciones puramente geométricas (excepto en su suposición de que se puede dibujar la parábola), el conocimiento de que z satisface una ecuación y las propiedades geométricas del círculo y la parábola (que se pueden ver más fácilmente mediante sus ecuaciones). Descartes hace aquí lo mismo que en el libro I, con la salvedad de que en el problema de construcciones que resuelve ahora la longitud incógnita satisface una ecuación de tercer grado o superior, en vez de ser de primer o segundo grados. Su solución de los aspectos puramente algebraicos del problema y la subsiguiente construcción son prácticamente las mismas que habían dado los árabes, con la diferencia de que él podía usar las ecuaciones de las secciones cónicas para deducir sus propiedades y dibujarlas.

Descartes no sólo deseaba mostrar cómo podían resolverse ciertos problemas sólidos con la ayuda del álgebra y de las cónicas, sino que le interesaba clasificar los problemas para poder entender lo que había en cada uno de ellos y saber cómo proceder para resolverlos. Su clasificación se basa en el grado de la ecuación a la que se llega al formular algebraicamente el problema de construcción. Si dicho grado es uno o dos, la construcción puede efectuarse con rectas y círculos. Si el grado es tres o cuatro, hay que usar cónicas, y afirma —por cierto— que todos los problemas cúbicos pueden reducirse a la trisección del ángulo y la duplicación del cubo, y que ninguno de ellos puede resolverse sin utilizar curvas más complicadas que el círculo. Si el grado de la ecuación es superior a cuatro, se precisan curvas más complicadas que las cónicas para efectuar la construcción.

Descartes también consideraba el grado de una curva como medida de su simplicidad. Siempre debe usarse la curva más simple, es decir, la de menor grado, para resolver un problema de construcción. Esta insistencia en el grado llegó a ser tan intensa que una figura complicada como el folium (u hoja) de Descartes (fig. 15.4), cuya ecuación es $x^3 + y^3 - 3axy = 0$, era considerada más simple que $y = x^4$.

Mucho más importante que la visión cartesiana de los problemas

de construcción y su clasificación es la importancia dada al álgebra. Esta es la clave para reconocer los problemas típicos de la geometría y unificar problemas cuya forma geométrica no parece guardar relación alguna. El álgebra aporta a la geometría los principios más naturales de clasificación y la jerarquía de método más natural. No sólo pueden discutirse con elegancia, plenitud y rapidez las cuestiones de resolubilidad y constructibilidad geométrica gracias al álgebra, sino que sin ella tal discusión es imposible. De esta forma, el sistema y la estructura fueron transferidos de la geometría al álgebra.

Parte del libro II de *La Géometrie* y toda *La Dioptrique* están dedicadas a la óptica, con la geometría analítica como auxiliar. A Descartes le interesaba mucho el diseño de lentes para el telescopio, el microscopio y otros instrumentos ópticos, pues era consciente de la gran importancia de estos instrumentos en astronomía y biología. En la *Dioptrique* considera el fenómeno de la refracción. Antes que él, Kepler y Alhazén habían observado que la hipótesis usual de proporcionalidad entre el ángulo de refracción y el de incidencia, con constante dependiente del medio refractor, es incorrecta para ángulos grandes, aunque no llegaron a descubrir la ley verdadera. Antes de 1626, Willebrord Snell había descubierto, pero no publicado, la relación correcta:

$$\frac{\text{sen } i}{\text{sen } r} = \frac{v_1}{v_2},$$

donde v_1 es la velocidad de la luz en el primer medio (fig. 15.5) y v_2 la correspondiente al segundo, al cual pasa la luz. Descartes dio la misma

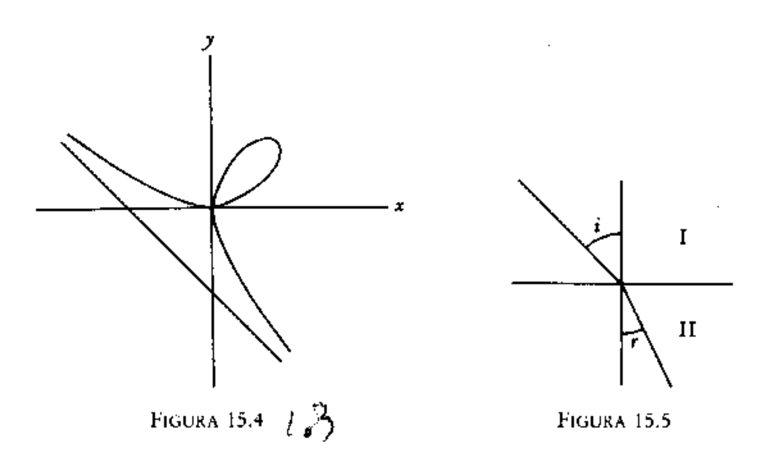


FIGURA 15.4

FIGURA 15.5

ley en 1637 en *La Dioptrique*, y hay cierta discusión con respecto a si la descubrió independientemente. Su argumento era incorrecto, e inmeditamente Fermat abordó tanto la propia ley como su demostración. Surgió entre ambos una controversia que duró diez años. Fermat no estuvo convencido de la validez de la ley hasta que la dedujo de su principio de tiempo mínimo (cap. 24, sec. 3).

En *La Dioptrique*, tras describir el funcionamiento del ojo, Descartes considera el problema de diseñar lentes para telescopios, microscopios y gafas. Se sabe desde la antigüedad que una lente esférica no enfoca en un solo punto los rayos paralelos o divergentes desde una fuente. Estaba, pues, abierta la cuestión de la forma que habría de tener una lente para enfocar los rayos incidentes. Kepler había sugerido que alguna cónica podría tener tal propiedad. Descartes se aplicó al diseño de una lente que enfocase los rayos perfectamente.

Procedió a resolver el problema general: ¿qué superficie de separación entre dos medios es tal que los rayos de luz procedentes de un punto del primer medio que pasen por refracción al segundo convergen en un punto? Descubrió que la curva generatriz de la superficie de revolución buscada es un óvalo, actualmente llamado óvalo de Descartes. Esta curva, junto con sus propiedades de refracción, se discuten en *La Dioptrique*, y esta discusión se complementa en el libro II de *La Géometrie.*

Según la definición moderna, el óvalo es el lugar geométrico de los puntos M que satisfacen la condición

$$FM \pm nF'M = 2a$$

donde F y F' son dos puntos fijos, $2a$ es un número real mayor que FF' y n es un número real arbitrario. Si $n = 1$, la curva es una elipse. En el caso general, la ecuación del óvalo es de cuarto grado en x e y, y la figura consta de dos curvas cerradas y distintas, una contenida en la otra y sin punto común. La curva interior es convexa y la exterior puede serlo también o tener puntos de inflexión como en la figura 15.6.

Vemos, pues, que la visión de Descartes de la geometría de coordenadas difiere profundamente de la de Fermat. Descartes criticaba la tradición griega y proponía romper con ella, mientras que Fermat creía en la continuidad con el pensamiento griego, y consideraba su trabajo en geometría analítica como una mera reformulación de la obra de Apolonio. El verdadero descubrimiento, la potencia de

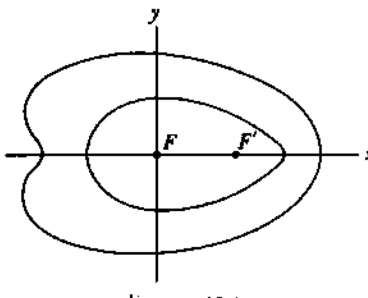


Figura 15.6

los métodos algebraicos, corresponde a Descartes, que sabía que su método estaba suplantando a los antiguos. Aunque la idea de asociar ecuaciones a las curvas es más clara en Fermat que en Descartes, el trabajo de aquél es fundamentalmente un logro técnico que completa la obra de Apolonio y explota la idea de Vieta de la representación lineal. El método de Descartes es de aplicación universal y potencialmente aplicable también a las curvas trascendentes.

A pesar de diferencias tan significativas en cuanto a método y objetivos de la geometría analítica, Descartes y Fermat se enredaron en una discusión acerca de la prioridad del descubrimiento. La obra de Fermat no se publicó hasta 1679; sin embargo, su descubrimiento de las ideas básicas de la geometría analítica, en 1629, antecede a la publicación de *La Géometrie* de Descartes en 1637. En esta época, Descartes conocía plenamente muchos hallazgos de Fermat, pero siempre negó que sacara sus ideas de él. Las ideas de Descartes en geometría analítica, según el matemático holandés Isaac Beeckman (1588-1637), se remontan a 1619. Además, no hay posible discusión acerca de la originalidad de muchas de sus ideas básicas sobre el tema.

Cuando se publicó *La Géometrie,* Fermat la criticó por omitir ideas como las de máximos y mínimos, las tangentes a las curvas y la construcción de lugares geométricos sólidos, temas que, en su opinión, merecían la atención de cualquier geómetra. Descartes, por su parte, dijo que Fermat había hecho muy poca cosa; de hecho, nada a lo que no pudiera llegarse fácilmente con trabajo o conocimiento previo; mientras que él había necesitado una profunda experiencia sobre la naturaleza de las ecuaciones, tal y como aparece en el tercer libro de *La Géometrie.* Descartes se refería sarcásticamente a Fermat como *vostre Conseiller de Maximis et Minimis,* y afirmaba que estaba

en deuda con él. Roberval, Pascal y otros tomaron el partido de Fermat, y Mydorge y Desargues, el de Descartes. Los amigos de Fermat escribieron implacables cartas contra Descartes. Con el tiempo, las actitudes de ambos se suavizaron, y en un trabajo de 1660, Fermat, mientras apuntaba cierto error en *La Géometrie*, declaró que admiraba de tal forma su genio, que, incluso cuando cometía errores, el trabajo de Descartes era más valioso que el de otros que no cometían ninguno. Descartes no había sido tan generoso con él.

El acento que la posteridad ha puesto sobre *La Géometrie* no es el que interesaba a Descartes. Aunque la idea sobresaliente para el futuro de las matemáticas era la de asociar ecuación y curva, para Descartes esto no era más que un medio para un fin, a saber, la resolución de problemas de construcciones geométricas. El énfasis de Fermat en las ecuaciones de lugares geométricos es, desde el punto de vista moderno, más oportuno. Las construcciones geométricas que con tanto ahínco describe Descartes en los libros I y III han ido perdiendo su importancia, en buena medida porque la construcción ya no se usa, como entre los griegos, para establecer la existencia del objeto construido.

Hay otra parte del libro III que también ha encontrado un lugar permanente en la matemática. Dado que Descartes resolvía los problemas de construcciones geométricas formulándolos primero algebraicamente, resolviendo después las ecuaciones obtenidas, y efectuando finalmente las construcciones que exigiese la solución, con el tiempo llegó a reunir, junto con sus propias investigaciones en la teoría de ecuaciones, otras que habían facilitado la resolución de algunos de sus problemas. Y dado que las ecuaciones algebraicas siguieron apareciendo en cientos de contextos distintos que nada tenían que ver con construcciones geométricas, esta teoría de ecuaciones se ha convertido en parte básica del álgebra elemental.

5. El avance de la geometría analítica durante el siglo XVII

La idea fundamental de la geometría analítica, el empleo de ecuaciones algebraicas para representar y estudiar curvas, no fue acogida con entusiasmo por los matemáticos por varias razones. El libro de Fermat, *Ad Locos*, aunque circuló entre sus amigos, no fue publicado hasta 1679. El interés que puso Descartes en resolver problemas de construcciones geométricas oscureció la idea principal,

la de ecuación asociada a una curva. De hecho, muchos de sus contemporáneos consideraban la geometría analítica como una herramienta para resolver los problemas de construcción. Leibniz llegó a referirse a la obra de Descartes como una regresión a los antiguos. El propio Descartes se dio cuenta de que su aportación iba mucho más allá de contribuir con un nuevo método de resolución de problemas de construcciones. En la introducción a *La Géometrie* dice: «Además, lo que he escrito en el segundo libro sobre la naturaleza y propiedades de las líneas curvas y sobre el método de examinarlas va, a mi entender, mucho más allá del tratamiento de la geometría ordinaria, como la retórica de Cicerón va más allá del *a*, *b*, *c* de los niños.» Sin embargo, los usos alternativos que dio a las ecuaciones de las curvas, como la resolución del problema de Pappus, la obtención de normales a las curvas o el estudio de las propiedades de los óvalos, se vieron oscurecidos por la atención que dedicó a los problemas de construcción. Otra razón de la lenta difusión de la geometría analítica fue la insistencia de Descartes en hacer su presentación tan difícil de seguir.

Además de esto, muchos matemáticos ponían objeciones a la idea de confundir álgebra y geometría, o aritmética y geometría. Estas protestas se venían oyendo ya desde el siglo XVI, durante el ascenso del álgebra. Tartaglia, por ejemplo, insistía en la distinción entre las operaciones con objetos geométricos a la manera griega y las operaciones entre números. Reprochó a un traductor de Euclides el emplear como sinónimos *multiplicare* y *ducere*. La primera pertenece al número, decía, y la segunda a la magnitud. También Vieta consideraba las ciencias del número y de la magnitud geométrica como paralelas pero distintas. El propio Newton, en su *Arithmetica Universalis*, no estaba de acuerdo en confundir álgebra y geometría, aunque hizo ciertas contribuciones a la geometría analítica y la empleó en el cálculo. Dice [8]:

Las ecuaciones son expresiones de cálculo aritmético y no tienen propiamente lugar en la geometría, excepto en el sentido de que permitan probar que cantidades geométricas verdaderas (esto es, líneas, superficies, sólidos y proporciones) coincidan unas con otras. Multiplicaciones, divisiones y cálculos de este tipo han sido recientemente introducidos en la geometría de forma poco aconsejable y contra los principios primeros de dicha ciencia... Por consiguiente, no deben confundirse ambas ciencias, y el fruto reciente de tal

[8] *Arithmetica Universalis*, 1707, p. 282.

confusión es la pérdida de esa simplicidad en la que consiste la elegancia geométrica.

Una interpretación razonable de la postura de Newton es que quería mantener el álgebra aparte de la geometría elemental, pero la consideraba útil en el estudio de las cónicas y curvas de grado superior.

Otra razón más de la lentitud de la aceptación de la geometría analítica era la crítica general a la falta de rigor del álgebra. Ya hemos mencionado la repugnancia de Barrow a aceptar los números irracionales como algo más que meros símbolos de magnitudes geométricas continuas (cap. 13, sec. 2). La aritmética y el álgebra hallaban su justificación lógica en la geometría; por tanto, el álgebra no podía reemplazar a la geometría o existir como su igual. El filósofo Thomas Hobbes (1588-1679), figura de pequeña talla en matemáticas, hablaba, sin embargo, en nombre de muchos cuando criticaba al «rebaño de los que aplican su álgebra a la geometría». Hobbes decía que esos algebristas consideraban erróneamente los símbolos como geometría, y describía el libro de John Wallis sobre cónicas como ruin y como «una costra de símbolos».

A pesar de todos estos obstáculos contra la justa apreciación de las contribuciones de Descartes y Fermat, un grupo de matemáticos fue progresivamente adoptando y extendiendo la geometría analítica. La primera tarea que había que realizar era explicar la idea de Descartes. Una traducción latina de *La Géometrie* de Frans van Schooten (1615-60), publicada por primera vez en 1649 y vuelta a editar muchas veces, no sólo puso el libro al alcance de todos los estudiosos en una lengua que podían leer, sino que contenía un comentario explicativo de la concentrada presentación de Descartes. En la edición de 1659-61, van Schooten daba efectivamente la forma algebraica de una transformación de coordenadas de una línea de base (eje de las x) en otra. Tan impresionado estaba del poder del método de Descartes que pretendía que los geómetras griegos lo habían empleado para deducir sus resultados. Disponiendo del material algebraico, los griegos, según van Schooten, vieron cómo obtener los resultados sintéticamente (y mostró cómo hacerlo) y estos métodos sintéticos, menos transparentes que los algebraicos, fueron los que publicaron para asombrar al mundo. Van Schooten puede haberse engañado por la palabra «análisis», que para los griegos significaba analizar un problema, y el término «geometría analítica», que describe específicamente el uso del álgebra como método por parte de Descartes.

John Wallis, en *De Sectionibus Conicis* (1655), dedujo primeramente las ecuaciones de las cónicas traduciendo las condiciones geométricas de Apolonio a la forma algebraica (de forma muy parecida a lo que hemos visto en el cap. 4, sec. 12), con objeto de discutir los resultados de aquél. Definió después las cónicas como las curvas correspondientes a ecuaciones de segundo grado en x e y, y probó que éstas eran efectivamente las secciones cónicas usuales de la geometría. Probablemente fue el primero en probar propiedades de las cónicas mediante sus ecuaciones. Su libro ayudó enormemente a difundir la idea de la geometría de coordenadas y a popularizar el tratamiento de las cónicas como curvas en el plano en vez de como secciones del cono, aunque esta versión tradicional siguió utilizándose. Wallis insistía además en la validez del razonamiento algebraico, mientras que Descartes, al menos en *La Géometrie*, se apoyaba realmente en la geometría, considerando el álgebra meramente como una herramienta. Wallis fue también el primero en introducir conscientemente abscisas y ordenadas negativas. Es posible que Newton, que hizo lo mismo más tarde, tomara la idea de Wallis. Podemos contrastar la observación de van Schooten acerca del método con la de Wallis, para quien Arquímedes y casi todos los clásicos habían ocultado a la posteridad su método de descubrimiento y análisis de forma tan completa que a los modernos les había resultado más fácil inventar un nuevo análisis que recuperar el antiguo.

El libro de Newton *The Method of Fluxions and Infinite Series* (El método de fluxiones y series infinitas), escrito hacia 1671, pero publicado por primera vez en traducción inglesa de John Colson (m. en 1760) con el mencionado título en 1736, contiene muchos usos de la geometría analítica, sobre todo el trazado de curvas a partir de sus ecuaciones. Una de las ideas originales que ofrece es el empleo de nuevos sistemas de coordenadas. Todos los matemáticos del siglo XVII y muchos del XVIII usaban generalmente un solo eje y trazaban los valores de y formando un ángulo recto u oblicuo con el eje, lo que corresponde en esencia a nuestro sistema de coordenadas polares. El libro contiene muchas variantes de esta idea. También introduce Newton las coordenadas bipolares, en las cuales se sitúa un punto de acuerdo con sus distancias a dos puntos fijos (fig. 15.7). Como esta obra de Newton no vio la luz hasta 1736, el mérito del descubrimiento de las coordenadas polares suele atribuírsele a Jacobo Bernoulli, que publicó un artículo al respecto en las *Acta Eruditorum* de 1691.

Se introdujeron muchas curvas nuevas, junto con sus ecuaciones.

En 1694, Bernoulli introdujo la lemniscata [9], que desempeñó un papel de gran importancia en todo el análisis del siglo XVIII. Se trata de un caso particular de una clase de curvas llamadas óvalos de Cassini (lemniscatas generales) introducidas por Jean-Dominique Cassini (1625-1712), aunque no aparecieron impresas hasta que su hijo Jacques (1677-1756) publicó los *Elements d'astronomie* en 1749. Los óvalos de Cassini (fig. 15.8) están definidos por la condición de que el producto $r_1 r_2$ de las distancias de un punto arbitrario de la curva a dos puntos fijos S_1 y S_2 sea igual a b^2, donde b es una constante. Sea $2a$ la distancia $S_1 S_2$. Si $b > a$, obtenemos un óvalo que no se corta a sí mismo; si $b = a$, resulta la lemniscata introducida por Jacobo Bernoulli, y si $b < a$ se tienen dos óvalos separados. La ecuación en coordenadas rectangulares de los óvalos de Cassini es de cuarto grado. El propio Descartes introdujo la espiral logarítmica [10], cuya ecuación en coordenadas polares es $\varrho = a^{\theta}$, y descubrió muchas de sus propiedades. Todavía aparecieron más curvas, como la catenaria y la cicloide, de las que hablaremos en relación con otras cuestiones.

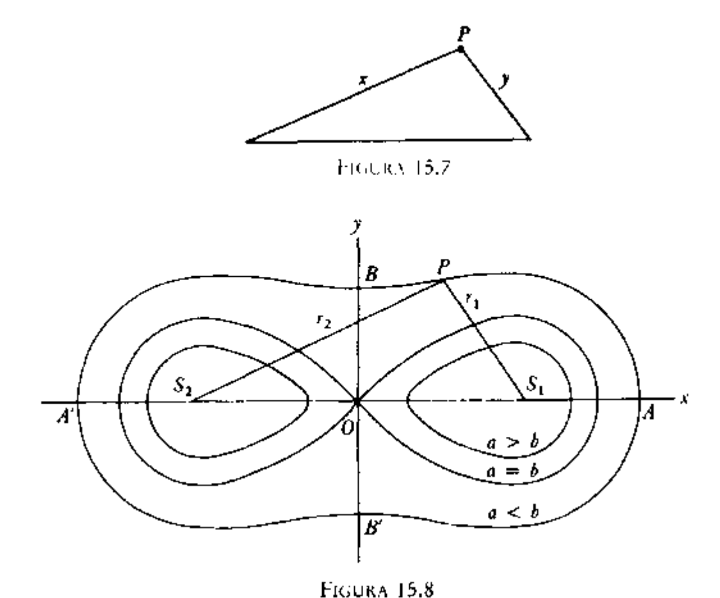


FIGURA 15.7

FIGURA 15.8

[9] *Acta Erud.*, sept. 1694 = *Opera*, 608-12.

[10] Carta a Mersenne del 12-IX-1638 = *Oeuvres*, 2, 360.

La extensión de la geometría analítica a tres dimensiones comenzó a elaborarse en el siglo XVII. En el libro II de *La Géometrie,* Descartes hace notar que sus ideas pueden aplicarse fácilmente a todas las curvas que puedan concebirse como engendradas por los movimientos regulares de un punto en el espacio tridimensional. Para representar algebraicamente una de tales curvas traza desde cada punto de la misma las perpendiculares a dos planos que se cortan en ángulo recto (fig. 15.9). Cada uno de los extremos de estas perpendiculares describe una curva en el plano respectivo, que puede tratarse por el método ya indicado. Mas al principio del libro II, Descartes observa que una ecuación en tres incógnitas que caracterice al punto general *C* de un lugar geométrico representa un plano, una esfera o una superficie más complicada. Evidentemente, se dio cuenta de que su método podía extenderse a curvas y superficies en el espacio tridimensional, pero no llegó a efectuar por sí mismo tal extensión.

Fermat da, en una carta de 1643, un breve esquema de sus ideas sobre la geometría analítica tridimensional. Habla de superficies cilíndricas, paraboloides elípticos, hiperboloides de dos hojas y elipsoides, y afirma que, para coronar la introducción a las curvas planas, habría que estudiar las curvas sobre superficies. «Esta teoría es susceptible de tratarse por un método general que explicaré si tengo tiempo.» En un trabajo de media página, *Novus Secundarum* [11], dice que una ecuación con tres incógnitas define una superficie.

La Hire, en sus *Nouveaux éléments des sections coniques* (1679), fue algo más preciso acerca de la geometría analítica tridimensional. Para representar una superficie, primero representa un punto *P* en el espacio mediante tres coordenadas, como se indica en la figura 15.10, y llega, de hecho, a escribir la ecuación de una superficie. Sin embargo, el desarrollo de la geometría analítica en tres dimensiones es obra del siglo XVIII, y hablaremos de ello más adelante.

6. La importancia de la geometría analítica

A la luz del considerable progreso que ya había realizado el álgebra antes de que Fermat y Descartes hicieran su aparición en la escena, la geometría de coordenadas no tiene la apariencia de un gran triunfo técnico. Para Fermat, se trataba de parafrasear algebraicamen-

---

[11] *Oeuvres*, 1, 186-87; 3, 161-62.

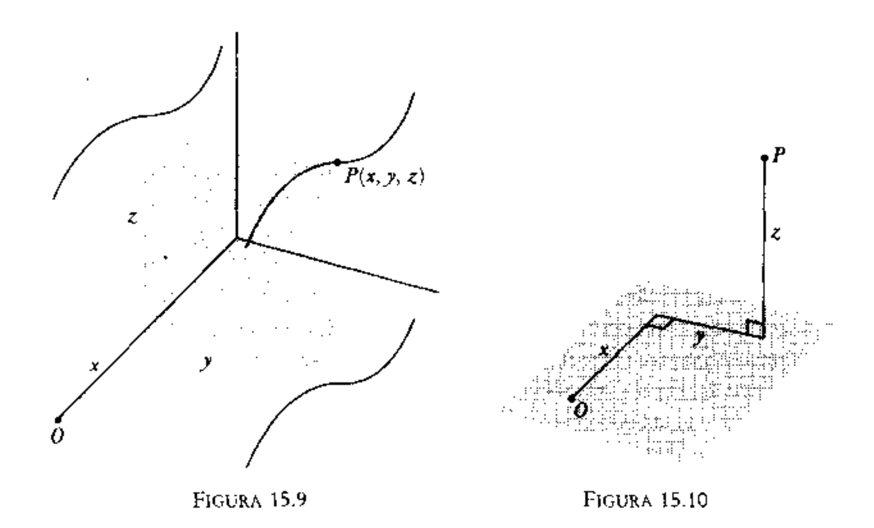


FIGURA 15.9

FIGURA 15.10

te a Apolonio. En el caso de Descartes, apareció casi por accidente mientras proseguía el trabajo de Vieta y otros empeñados en facilitar, mediante la introducción del álgebra, la resolución de ciertos problemas de construcciones geométricas. Y, sin embargo, la geometría analítica cambió de hecho la faz de las matemáticas.

Al argumentar que una curva es cualquier lugar geométrico que tiene una ecuación algebraica, Descartes abrió de un solo golpe el dominio matemático. Cuando se considera la variedad de curvas que han ido teniendo aceptación y uso en las matemáticas, y se compara tal conglomerado con lo que aceptaban los griegos, es fácil ver cuán importante fue el derribo de las barreras de la geometría griega.

Por medio de la geometría analítica, Descartes intentó dotar de método a la geometría, y consiguió mucho más de lo que nunca pensó. Es hoy un lugar común reconocer no sólo cuán fácilmente se puede probar muchas propiedades de las curvas con ayuda del álgebra, sino también que el método para abordarlas es casi automático. Más potente aún es el método: cuando Wallis y Newton empezaron a usar letras para designar tanto números positivos como negativos, llegando incluso después a referirse a números complejos, fue posible resumir en un solo tratamiento algebraico muchos casos que la geometría pura tenía que considerar separadamente. Por ejemplo, para demostrar, en geometría sintética, que las alturas de un triángulo

se cortan en un punto, hace falta considerar por separado si la intersección tiene lugar en el interior o en el exterior del triángulo. En geometría analítica, ambos casos se consideran conjuntamente.

La geometría analítica resultó ser una herramienta de doble uso para las matemáticas. Por una parte, los conceptos geométricos podían formularse algebraicamente, y los objetivos geométricos podían alcanzarse por medio del álgebra. Recíprocamente, al interpretar geométricamente los enunciados algebraicos puede lograrse una visión más intuitiva de su significado, lo cual puede, a su vez, ser fuente de nuevas conclusiones. Estas virtudes menciona Lagrange en sus *Leçons élémentaires sur les mathématiques* [12]: «Mientras álgebra y geometría llevaron caminos separados, su progreso fue lento y sus aplicaciones limitadas. Pero en cuanto ambas ciencias juntaron sus fuerzas, sacaron cada una de la otra una vitalidad fresca y desde entonces desfilaron a paso ligero hacia la perfección.» Ciertamente, la inmensa potencia que ha desarrollado la matemática desde el siglo XVII en adelante debe atribuirse, en gran medida, a la geometría analítica.

El mérito más importante de la geometría analítica fue dotar a la ciencia del utillaje matemático que siempre había necesitado, y que había empezado a exigir abiertamente en el siglo XVII: herramientas cuantitativas. No hay duda de que el estudio del mundo físico pide, en una primera etapa, geometría. Los objetos son básicamente figuras geométricas, y las trayectorias de los objetos móviles son curvas. El propio Descartes estaba, desde luego, convencido de que toda la física podía reducirse a geometría. Pero, como ya hemos apuntado, los usos de la ciencia en geodesia, navegación, cálculos de calendario, predicciones astronómicas, movimiento de proyectiles e incluso en el diseño de lentes, al que el mismo Descartes se dedicó, exigen conocimiento cuantitativo. La geometría analítica posibilitó la expresión de formas y trayectorias en forma algebraica, de la cual se podía extraer información cuantitativa.

Así pues, el álgebra, que Descartes había considerado como mera herramienta, más como una extensión de la lógica que como parte de las matemáticas propiamente dichas, llegó a ser más esencial que la propia geometría. La geometría analítica allanó de hecho el camino para la más completa inversión de papeles entre álgebra y geometría. Desde la época de los griegos hasta 1600, aproximadamente, la

---

[12] *Oeuvres*, 7, 183-287, p. 271 en parte.

geometría dominó las matemáticas y el álgebra estaba subordinada a ella; a partir de 1600, el álgebra se convirtió en la disciplina matemática básica, y en esta transformación el cálculo iba a ser el factor esencial. Los antecedentes del álgebra hicieron más difíciles las cosas por la causa que ya hemos mencionado: su falta de fundamentación lógica. Nada se haría al respecto, sin embargo, hasta bien entrado el siglo XIX.

El hecho de que el álgebra se edificase sobre una base empírica ha dado lugar a una confusión en la terminología matemática. La disciplina creada por Fermat y Descartes suele llamarse «geometría analítica». La palabra «analítica» no es apropiada; «geometría de coordenadas» o «geometría algebraica» (que tiene hoy día otro significado) hubiera sido más adecuado. La palabra «análisis» se ha venido empleando desde Platón para significar el proceso de remontarse partiendo de lo que se desea probar hasta llegar a alguna verdad conocida. Es, en este sentido, opuesto a «síntesis», que describe la presentación deductiva. Hacia 1590, Vieta rechazó la palabra «álgebra» por carecer de significado en las lenguas europeas, y propuso el término «análisis» (cap. 13, sec. 8), pero no se llegó a adoptar esta sugerencia. Sin embargo, para él y para Descartes, la palabra «análisis» era todavía apropiada para describir la aplicación del álgebra a la geometría, ya que el álgebra servía para «analizar» el problema de construcción geométrica considerado: se suponía conocida la longitud buscada, se hallaba una ecuación satisfecha por esa longitud, se operaba con dicha ecuación y se concluía con la construcción de la longitud en cuestión. Así, Jacques Ozanam (1640-1717) dice en su *Diccionario* (1690) que los geómetras modernos efectuaban sus análisis por medio del álgebra. En la famosa *Encyclopédie* del XVIII, d'Alembert usa «álgebra» y «análisis» como sinónimos. Gradualmente, «análisis» llegó a significar el método algebraico, aunque la nueva geometría de coordenadas, hasta el fin del siglo XVIII, era generalmente descrita como la aplicación del álgebra a la geometría. Para fines de siglo, el término «geometría analítica» llegó a ser el usual y a aparecer con frecuencia en los títulos de los textos.

Sin embargo, a medida que el álgebra se convertía en la disciplina dominante, los matemáticos empezaron a considerar que poseía una función mucho mayor que la de analizar un problema en el sentido de los griegos. En el siglo XVIII, el punto de vista de que el álgebra, en su aplicación a la geometría, es algo más que una herramienta, que la propia álgebra es un método de introducir y estudiar curvas y

superficies (el supuesto punto de vista de Fermat, en oposición al de Descartes) terminó por triunfar como resultado de la obra de Euler, Lagrange y Monge. De aquí que el término «geometría analítica» se refiera tanto a la demostración como al uso del método algebraico, y, en consecuencia, hablemos hoy de geometría analítica en oposición a geometría sintética, sin significar que una es un método de invención y la otra de demostración, pues ambas geometrías son deductivas.

Mientras tenían lugar estas transformaciones, el cálculo y algunas extensiones suyas, como las series infinitas, entraban en el campo matemático. Tanto Newton como Leibniz consideraban el cálculo como una extensión del álgebra; se trata del álgebra del infinito, o del álgebra que trata con cantidades infinitas de términos, como en las series infinitas. En fecha tan tardía como 1797, Lagrange, en su *Théorie des fonctions analytiques*, decía que el cálculo y sus desarrollos eran sólo generalizaciones del álgebra elemental. Como álgebra y análisis habían sido sinónimos, el propio cálculo llegó a ser llamado análisis. En un famoso texto de 1748, Euler usó el término «análisis infinitesimal» para describir el cálculo. Este término fue empleado hasta finales del siglo XIX, en que se adoptó la palabra «análisis» para describir el cálculo y las demás ramas de las matemáticas construidas sobre él. La situación resultante es, pues, confusa, pues el término «análisis» abarca todos los desarrollos basados en los límites, mientras que la «geometría analítica» no contiene procesos de límite.

Bibliografía

Boyer, Carl B.: *History of Analytic Geometry.* Scripta Mathematica, 1956.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1900, Johnson Reprint Corp., 1965, vol. 2, pp. 806-76.

Chasles, Michel: *Aperçu historique des méthodes en géometrie*, 3 ed., Gauthier-Villars et Fils, 1889, caps. 2-3 y notas.

Coolidge, Julian L.: *A History of Geometrical Methods.* Dover (reimpresión), 1963, pp. 117-31.

Descartes, René: *The Geometry*, Dover (reimpresión), 1954.

—: *Oeuvres*, 12 vols., Cerf, 1897-1913.

—: *El discurso del método*, Madrid, Alfaguara, 1987.

Fermat, Pierre de: *Oeuvres*, 4 vols. y suplemento, Gauthier-Villars, 1891-1912; suplemento, 1922.

Montucla, J. F.: *Histoire des mathématiques*, Albert Blanchard (reimpresión), 1960, vol. 2, pp. 102-77.

Scott, J. F.: *The Scientific Work of René Descartes,* Taylor and Francis, 1952.
Smith, David E.: *A Source Book in Mathematics,* 2 vols., Dover (reimpresión), 1959, vol. 2, pp. 389-402.
Struik, D. J.: *A Source Book in Mathematics, 1200-1800,* Harvard University Press, 1969, pp. 87-93, 143-157.
Wallis, John: *Opera,* 3 vols. (1693-99), Georg Olms (reimpresión), 1968.
Vrooman, Jack, R.: *René Descartes: A Biography,* G. P. Putnam's Sons, 1970.
Vuillemin, Jules: *Mathématiques et métaphysique chez Descartes*, Presses Universitaires de France, 1960.

Capítulo 16
LA MATEMATIZACION DE LA CIENCIA

> De manera que podemos decir que la puerta está ahora abierta, por primera vez, a un método nuevo, acompañado por numerosos y maravillosos resultados que, en años venideros, atraerá la atención de otras mentes.
>
> GALILEO GALILEI

1. Introducción

Hacia 1600, los científicos europeos estaban indudablemente impresionados por la importancia de las matemáticas en el estudio de la naturaleza. La prueba más fuerte de esta convicción era el deseo de Copérnico y Kepler de derrocar las leyes aceptadas de la astronomía y de la mecánica, y las doctrinas religiosas al respecto, por una teoría que tenía sólo ventajas matemáticas. Sin embargo, el éxito sorprendente de la ciencia moderna y el impulso enorme para el trabajo creativo que las matemáticas encontraron en esta fuente no se hubieran producido si la ciencia hubiera continuado como en el pasado. Pero en el siglo XVII, dos hombres, Descartes y Galileo, revolucionaron la misma naturaleza de la actividad científica. Seleccionaron los conceptos que debía utilizar la ciencia, redefinieron los objetivos de la actividad científica y alteraron el método de la ciencia. Su reformulación no sólo suministró una potencia inesperada y sin precedentes a la ciencia, sino que la ligó indisolublemente a las matemáticas. Para entender el espíritu que animó a las matemáticas desde el siglo XVII hasta el XIX, debemos examinar primero las ideas de Descartes y de Galileo.

2. El concepto de la ciencia de Descartes

Descartes proclamó explícitamente que la esencia de la ciencia eran las matemáticas. Dice que «ni admite ni espera ningún principio de la física diferente de los que están en la geometría o en la matemática abstracta, porque así se explican todos los fenómenos de la naturaleza y pueden darse algunas demostraciones de ellos». El mundo objetivo es espacio solidificado, o geometría encarnada. Sus propiedades deben poderse deducir, por tanto, del primer principio de la geometría.

Descartes explicó con más detalles que el mundo debe ser accesible y reducible a las matemáticas. Insistió en que las propiedades más fundamentales y fiables de la materia son forma, extensión y movimiento en el espacio y en el tiempo. Como la forma es sólo extensión, Descartes afirmaba: «Dadme extensión y movimiento y construiré el universo.» El movimiento en sí mismo se producía por la acción de las fuerzas sobre las moléculas. Descartes estaba convencido de que estas fuerzas obedecían a leyes matemáticas invariables y, puesto que la extensión y el movimiento eran expresables matemáticamente, todos los fenómenos podían ser descritos matemáticamente.

La filosofía mecanicista de Descartes se extendía incluso al funcionamiento del cuerpo humano. Creía que las leyes de la mecánica explicarían la vida del hombre y de los animales, y en sus trabajos de fisiología utilizó el calor, la hidráulica, tubos, válvulas y las acciones mecánicas de las palancas para explicar las acciones del cuerpo. Sin embargo, Dios y el alma estaban exentos de mecanismos.

Si Descartes consideraba el mundo externo como consistente sólo de materia en movimiento, ¿cómo explicaba los gustos, olores, colores y las cualidades de los sonidos? Aquí adoptaba la antigua doctrina griega de las cualidades primarias y secundarias la cual, según la exponía Demócrito, mantenía que «lo dulce y amargo, lo frío y caliente, así como los colores, todas estas cosas existen, pero en la opinión y no en la realidad; lo que realmente existe son partículas inmutables, átomos, y sus movimientos en el espacio vacío». Las cualidades primarias, materia y movimiento, existen en el mundo físico; las cualidades secundarias son sólo efectos que las cualidades primarias inducen en los órganos de los sentidos de los seres humanos por el impacto de átomos externos en esos órganos.

Por lo tanto, para Descartes hay dos mundos; uno, una enorme máquina matemática, armoniosamente diseñada, que existe en el

espacio y en el tiempo, y otro, el mundo de los espíritus pensantes. El efecto de los elementos del primer mundo en el segundo produce las cualidades no matemáticas o secundarias de la materia. Descartes afirmaba, además, que las leyes de la naturaleza son invariables, porque son sólo una parte de un modelo matemático predeterminado, y que Dios no podía alterar la invariable naturaleza. Aquí Descartes negaba la extendida creencia de que Dios intervenía continuamente en el funcionamiento del universo.

Aunque las doctrinas filosóficas y científicas de Descartes subvertían el aristotelismo y el escolasticismo, él era escolástico en un aspecto fundamental: extrajo de su propia mente proposiciones acerca de la naturaleza del ser y de la realidad. Creía que existen verdades a priori y que el intelecto, por su propio poder, puede llegar a un conocimiento perfecto de todas las cosas; estableció, por ejemplo, leyes del movimiento sobre la base de un razonamiento a priori. (En realidad, en sus trabajos biológicos realizó experimentos, y extrajo de ellos conclusiones vitales.) Sin embargo, aparte de su confianza en los principios a priori, promulgó una filosofía sistemática y general que destrozó el soporte del escolasticismo y abrió frescos canales de pensamiento. Su intento de barrer todas las preconcepciones y los prejuicios era una clara declaración de revuelta con respecto al pasado. Al reducir los fenómenos naturales a acontecimientos puramente físicos, hizo mucho por desembarazar a la ciencia del misticismo y de las fuerzas ocultas. Los escritos de Descartes fueron altamente influyentes; su filosofía deductiva y sistemática se extendió en el siglo XVII e impresionó a Newton, especialmente por la importancia del movimiento en ella. Exposiciones de su filosofía delicadamente encuadernadas adornaban incluso los tocadores de las damas.

3. El enfoque de la ciencia de Galileo

Aunque la filosofía de la ciencia de Galileo Galilei coincide en gran parte con la de Descartes, fue Galileo quien formuló los más radicales, efectivos y concretos métodos de la ciencia moderna y quien, mediante su propio trabajo, demostró su efectividad.

Galileo (1564-1642) nació en Pisa en la familia de un comerciante en tejidos, e ingresó en la universidad de Pisa para estudiar medicina. Los cursos estaban allí todavía, aproximadamente, al nivel del *curriculum* medieval; Galileo aprendió sus matemáticas privadamente, de un

ingeniero práctico, y a la edad de diecisiete años cambió de la medicina a las matemáticas. Después de cerca de ocho años de estudio solicitó una plaza para enseñar en la universidad de Bolonia, pero fue rechazado por no tener méritos suficientes. Se aseguró una plaza de profesor en Pisa. Estando allí, comenzó a atacar a la ciencia aristotélica, y no dudó en expresar sus puntos de vista, aunque sus críticas le enemistaron con sus colegas. Había empezado también a escribir importantes trabajos matemáticos que suscitaron celos en los menos competentes. Se intentó, pues, que no se encontrara a gusto, y en 1592 se marchó y aceptó un puesto como profesor de matemáticas en la universidad de Padua. Allí escribió un libro corto, *Le mecaniche* (1604). Después de dieciocho años en Padua fue invitado a Florencia por el Gran Duque Cósimo II de Médicis, quien le nombró Matemático Principal de su corte, le dio casa y un salario considerable, y le protegió de los jesuitas, quienes dominaban el Papado y ya le habían amenazado por defender la teoría de Copérnico. Para expresar su gratitud, Galileo llamó a los satélites de Júpiter, que había descubierto durante el primer año al servicio de Cósimo, las estrellas Mediceas. En Florencia, Galileo tuvo la tranquilidad suficiente para proseguir sus estudios y para escribir.

Su defensa de la teoría de Copérnico molestó a la Inquisición romana, y en 1616 fue llamado a Roma. Sus enseñanzas sobre la teoría heliocéntrica fueron condenadas por la Inquisición; tuvo que prometer no publicar nada más sobre el tema. En 1630, el Papa Urbano VIII le dio permiso para publicar siempre que hiciera un libro matemático y no doctrinal. Por tanto, en 1632, publicó su clásico *Dialogo dei massimi sistemi* (Diálogo sobre el gran sistema del mundo). La Inquisición romana le llamó otra vez en 1633 y, bajo la amenaza de tortura, le obligaron a retractarse de su defensa de la teoría heliocéntrica. Se le prohibió publicar y se le exigió que viviera prácticamente bajo arresto domiciliario. Emprendió la tarea de escribir sus reflexiones de esos años y trabajó sobre los fenómenos del movimiento y sobre la resistencia de materiales. El manuscrito, titulado *Discorsi e dimostrazioni matematiche intorno à due nuove scienze (Discursos y demostraciones matemáticas referentes a dos nuevas ciencias,* que también se conoce como *Diálogos referentes a dos nuevas ciencias),* fue llevado secretamente a Holanda y publicado allí en 1638. Esta es la obra, clásica, en la que Galileo ofreció su nuevo método científico. Defendió sus acciones con las palabras de que nunca había «declinado en piedad y reverencia por la Iglesia y mi propia conciencia».

Galileo fue un hombre extraordinario en muchos campos. Fue un perspicaz observador astronómico. Se le llama a menudo el padre de la invención moderna; aunque no inventó el telescopio o los «vidrios asombrosos», como los llamaba Ben Jonson, fue capaz inmediatamente de construirlos cuando conoció la idea. Fue un inventor independiente del microscopio, y diseñó el primer reloj de péndulo. También diseñó un compás con escalas que proporcionaba automáticamente los resultados de cálculos numéricos, de modo que el usuario podía leer las escalas y evitar el tener que hacer los cálculos. Este aparato fue tan solicitado que produjo muchos para venderlos.

Galileo fue el primer estudioso moderno importante del sonido. Sugirió una teoría ondulatoria del sonido y comenzó a trabajar sobre el tono, los armónicos y la vibración de cuerdas. Su trabajo fue continuado por Mersenne y Newton y se convirtió en una inspiración importante para el trabajo matemático del siglo XVIII.

Los principales trabajos de Galileo, aunque referidos a temas científicos, se consideran todavía como obras maestras desde un punto de vista literario. Su *Sidereus Nuncius* (Mensajero sideral) de 1610, en el que anunciaba sus observaciones astronómicas y declaraba su apoyo a la teoría de Copérnico, fue un éxito inmediato, y fue elegido para la prestigiosa Academia dei Lincei en Roma. Sus dos más grandes clásicos, *Diálogo sobre los máximos sistemas* y los *Diálogos sobre dos nuevas ciencias*, son claros, directos, agudos, pero profundos. En ambos, Galileo hace que uno de los personajes presente los puntos de vista de la época, contra los que otro personaje razona hábil y tenazmente para mostrar las falacias y puntos débiles de estos esquemas, y la fuerza de los nuevos.

En su filosofía de la ciencia, Galileo rompió radicalmente con lo especulativo y lo místico en favor de una visión de la naturaleza mecánica y matemática. También creía que los problemas científicos no debían enredarse ni oscurecerse con argumentos teológicos. De hecho, uno de sus logros en la ciencia, aunque algo apartado del método que examinaremos a continuación, fue el de reconocer el terreno de la ciencia y separarlo del de las doctrinas religiosas.

Galileo, como Descartes, estaba convencido de que la naturaleza estaba diseñada matemáticamente. Su afirmación de 1610 es famosa:

La filosofía [naturaleza] está escrita en ese gran libro que siempre está delante de nuestros ojos —quiero significar el universo— pero que no podemos entender si no aprendemos primero el lenguaje, y comprendemos los símbo-

los, en los que está escrito. El libro está escrito en lenguaje matemático, y los símbolos son triángulos, circunferencias y otras figuras geométricas, sin cuya ayuda es imposible comprender ni una palabra de él, sin lo cual se deambula en vano a través de un oscuro laberinto [1].

La naturaleza es sencilla y ordenada, y su comportamiento es regular y necesario. Actúa de acuerdo con leyes matemáticas perfectas e inmutables. La razón divina es la fuente de lo racional en la naturaleza; Dios colocó en el universo esa rigurosa necesidad matemática que los hombres alcanzan sólo laboriosamente. El conocimiento matemático es, en consecuencia, no sólo verdad absoluta, sino tan sacrosanto como cualquier línea de las Sagradas Escrituras. De hecho, es superior, porque hay mucho desacuerdo con respecto a las Escrituras, pero no puede haber ninguna con respecto a las verdades matemáticas.

Otra doctrina, el atomismo del griego Demócrito, es más clara en Galileo que en Descartes. El atomismo presuponía un espacio vacío (que Descartes no aceptaba) y átomos indestructibles e individuales. El cambio consistía en la combinación y separación de los átomos. Todas las variedades cualitativas de los cuerpos eran debidas a la variedad cuantitativa en el número, tamaño, forma y disposición especial de los átomos. Las principales propiedades de los átomos eran la impenetrabilidad y la indestructibilidad; estas propiedades servían para explicar los fenómenos físicos y químicos. La adhesión de Galileo al atomismo le situó en la vanguardia de las doctrinas científicas.

El atomismo condujo a Galileo a la doctrina de las cualidades primarias y secundarias. Dice: «Si las orejas, lenguas y narices se suprimieran, soy de la opinión de que la forma, cantidad (tamaño) y movimiento permanecerían, pero se terminarían los olores, sabores y sonidos, los cuales, abstraídos de la criatura viviente, sólo son palabras.» De un plumazo, pues, tanto Galileo como Descartes simplificaron mil fenómenos y cualidades para concentrarse en la materia y el movimiento, propiedades que pueden ser descritas matemáticamente. Quizá no es muy sorprendente que en el siglo en el que los problemas de movimiento eran los más prominentes y serios, los científicos encontraran que el movimiento era un fenómeno físico fundamental.

La concentración en la materia y el movimiento fue sólo el primer

---

[1] *Opere*, 4. 171.

paso en el nuevo enfoque de la naturaleza de Galileo. Su siguiente pensamiento, también expresado por Descartes, fue que cualquier rama de la ciencia puede ser configurada sobre el modelo de las matemáticas. Esto implica dos pasos esenciales. Las matemáticas comienzan con axiomas —verdades claras y autoevidentes— y a partir de ellos se pasa a establecer nuevas verdades mediante razonamientos deductivos. Cualquier rama de la ciencia, pues, debe comenzar con axiomas o principios y continuar deductivamente. Además, se debe extraer de los axiomas tantas consecuencias como sea posible. Esta forma de pensar se remonta a Aristóteles, que también buscaba una estructura deductiva para la ciencia, con el modelo matemático *in mente*.

Sin embargo, Galileo se separó radicalmente de los griegos, de los científicos medievales, e incluso de Descartes en su método para obtener primeros principios. Los pregalileanos y Descartes habían creído que la mente proporcionaría los principios básicos; no había más que pensar en cualquier clase de fenómenos y la mente reconocería inmediatamente las verdades fundamentales. Este poder de la mente se ponía claramente de manifiesto en las matemáticas. Axiomas como «iguales sumados a iguales proporcionan iguales» o «dos puntos determinan una recta», sugerían inmediatamente el pensar en números o figuras geométricas, y eran verdades indudables. Así habían encontrado también los griegos algunos principios físicos igualmente atrayentes. Era apropiado el que todos los objetos del universo tengan un sitio natural. El estado de reposo parecía claramente más natural que el estado de movimiento. Parecía indudable, también, que tuviera que aplicarse una fuerza para poner y mantener a los cuerpos en movimiento. Creer que la mente proporciona principios fundamentales no se opone a que las observaciones pudieran jugar un papel en la obtención de esos principios. Pero las observaciones sólo evocaban los principios correctos, de la misma manera que la visión de una cara familiar puede traer a la mente hechos sobre esa persona.

Los científicos griegos y medievales estaban tan convencidos de que existían principios fundamentales a priori que cuando observaciones experimentales ocasionales no concordaban con ellos, inventaban explicaciones especiales para preservar los principios y explicar, aun así, las anomalías. Estos hombres, como decía Galileo, primero decidían cómo debía funcionar el mundo y luego adaptaban lo que veían a sus principios preconcebidos.

Galileo decidió que en física, en contraposición a lo que ocurría en matemáticas, los primeros principios deben proceder de la experiencia y de la experimentación. El camino para obtener principios básicos y correctos es prestar atención a lo que dice la naturaleza más que a lo que prefiere la mente. La naturaleza, razonaba, no hace primero el cerebro de los hombres y organiza a continuación el universo de manera que sea aceptable al intelecto humano. A los pensadores medievales, que continuaban repitiendo a Aristóteles y debatiendo lo que había querido decir, Galileo les dirigió la crítica de que el conocimiento proviene de la observación y no de los libros, y que era inútil debatir sobre Aristóteles. Dice: «Cuando disponemos de los decretos de la naturaleza, la autoridad no sirve para nada...» Por supuesto, también algunos pensadores del Renacimiento, y el contemporáneo de Galileo, Francis Bacon, habían llegado a la conclusión de que la experimentación era necesaria; en este aspecto particular de su nuevo método Galileo no estaba por delante de los demás. Sin embargo, el modernista Descartes no vio la sabiduría de la confianza de Galileo en la experimentación. Los hechos que provienen de los sentidos, decía Descartes, sólo pueden conducir al engaño, pero la razón penetra a través de esos engaños. Desde los principios generales innatos que proporciona la mente podemos deducir fenómenos particulares de la naturaleza y comprenderlos. Galileo también era consciente de que se podía obtener principios incorrectos mediante la experimentación y que, como consecuencia, las deducciones que se basaran en ellos podían ser incorrectas. Por ello propuso el empleo de los experimentos para comprobar las conclusiones de sus razonamientos así como para obtener principios básicos.

Galileo era, en realidad, una figura de transición por lo que se refiere a la experimentación. Él, con Isaac Newton cincuenta años más tarde, creía que unos pocos experimentos claves o críticos podrían proporcionar principios fundamentales correctos. Además, muchos de los llamados experimentos de Galileo eran en realidad experimentos mentales, es decir, confiaba en la experiencia común para imaginar lo que ocurriría si se realizara un experimento. Entonces él obtenía una conclusión con la misma confianza que si hubiera realizado el experimento. Cuando en el *Diálogo sobre el gran sistema del mundo* describe el movimiento de una bola arrojada desde el mástil de un barco en movimiento, uno de los personajes, Simplicio, le pregunta si había realizado un experimento, Galileo le contesta: «No, y no lo necesito, porque sin ninguna experiencia puedo confir-

mar que es así, ya que no puede ser de otra manera.» Dice, de hecho, que experimentaba pocas veces, y sobre todo, en esos casos, para refutar a quienes no se interesaban por las matemáticas. Aunque Newton realizó algunos famosos e ingeniosos experimentos, también dice que los utilizaba para hacer sus *resultados* físicamente inteligibles y para convencer a la gente.

La verdad del asunto es que Galileo tenía algunas ideas preconcebidas acerca de la naturaleza, lo que le hacía confiar en que unos pocos experimentos bastarían. Creía, por ejemplo, que la naturaleza era sencilla. Por tanto, cuando consideraba cuerpos que caían libremente, los cuales caen con velocidad creciente, suponían que el aumento de la velocidad es el mismo en cada segundo de caída. Esta era la «verdad» más sencilla. Creía también que la naturaleza está diseñada matemáticamente, y, por tanto, cualquier ley matemática que pareciera estar de acuerdo con ella, aun sobre la base de una experimentación bastante limitada, le parecía correcta.

Para Galileo, así como para Huygens y Newton, la parte matemática, deductiva, de la empresa científica tenía una importancia mayor que la experimental. Galileo no estaba menos orgulloso de la abundancia de teoremas que surgían de un único principio que del descubrimiento del principio mismo. Los hombres que forjaron la ciencia moderna —Descartes, Galileo, Huygens y Newton (podemos también incluir a Copérnico y a Kepler)— enfocaron el estudio de la naturaleza como matemáticos, en su método general y en sus investigaciones concretas. Fueron primordialmente pensadores especulativos que esperaban aprehender principios matemáticos amplios, profundos (pero también sencillos), claros e inmutables, bien a través de la intuición o mediante observaciones y experimentos cruciales, y deducir entonces nuevas leyes a partir de estas verdades fundamentales, enteramente en la forma en que las propias matemáticas habían construido su geometría. El grueso de la actividad era la porción deductiva; así se derivarían sistemas completos de pensamiento.

Lo que los grandes pensadores del siglo XVII consideraban como el método adecuado de la ciencia se reveló en realidad como el camino provechoso. La búsqueda racional de las leyes de la naturaleza produjo, en la época de Newton, resultados extremadamente valiosos sobre la base del conocimiento observacional y experimental más ligero. Los grandes avances científicos de los siglos XVI y XVII se hicieron en astronomía, donde la observación ofrecía poca cosa nueva, y en mecánica, en la que los resultados experimentales no eran

muy sorprendentes y ciertamente no decisivos, mientras que la teoría matemática correspondiente alcanzaba comprehensión y perfección. Y durante los dos siglos siguientes los científicos produjeron leyes de la naturaleza profundas y extensas, apoyados en muy pocos experimentos y en observaciones casi triviales.

La esperanza de Galileo, Huygens y Newton de que sólo unos pocos experimentos bastarían para sus propósitos, puede entenderse fácilmente. Como estaban convencidos de que la naturaleza estaba diseñada matemáticamente, no veían ninguna razón por lo que no pudieran actuar en asuntos científicos como los matemáticos actuaban en su campo. Como dice John Herman Randall en *Making of the Modern Mind:* «La ciencia nació de la fe en la interpretación matemática de la naturaleza...»

Galileo, sin embargo, obtuvo unos pocos principios a partir de la experiencia, y en este trabajo también su enfoque fue una ruptura radical con respecto al de sus predecesores. Pensaba que se debe penetrar en lo que es fundamental en los fenómenos y comenzar allí. En *Dos nuevas ciencias* dice que no es posible tratar la variedad infinita de pesos, formas y velocidades. Había observado que las velocidades con las que caen objetos disímiles difieren menos en el aire que en el agua. Por lo tanto, cuanto más ligero sea el medio hay menos diferencias entre las velocidades de caída de los cuerpos. «Habiendo observado esto, vino a mí la conclusión de que en un medio totalmente desprovisto de resistencia todos los cuerpos caerían con la misma velocidad.» Lo que Galileo estaba haciendo aquí era separar los efectos incidentales o sin importancia en un esfuerzo por llegar al principal.

Por supuesto que los cuerpos reales caen en un medio resistente. ¿Qué podía decir Galileo acerca de esos movimientos? Su respuesta fue: «... por tanto, para manejar este asunto de una forma científica, es necesario prescindir de estas dificultades (resistencia del aire, fricción, etcétera) y, habiendo descubierto y demostrado los teoremas en el caso en que no hay resistencia, utilizarlos y aplicarlos con las limitaciones que nos muestre la experiencia».

Habiendo prescindido de la resistencia del aire y de la fricción, Galileo buscaba leyes básicas del movimiento en el vacío. Por tanto, no sólo contradecía a Aristóteles e incluso a Descartes por pensar en cuerpos moviéndose en un espacio vacío, sino que hacía precisamente lo que hace el matemático al estudiar figuras reales. El matemático prescinde de la estructura molecular, del color y del grosor de las

líneas, para llegar a algunas propiedades básicas y concentrarse en ellas. Así se introdujo Galileo en los factores físicos básicos. El método matemático de la abstracción es, en verdad, un paso más allá de la realidad pero, paradójicamente, conduce de vuelta a la realidad con una fuerza mayor que si todos los factores realmente presentes se tienen en cuenta de una vez.

Hasta entonces, Galileo había formulado unos cuantos principios metodológicos, muchos de los cuales fueron sugeridos por el camino que las matemáticas habían utilizado en la geometría. Su siguiente principio fue el de utilizar las matemáticas mismas, pero en una forma especial. A diferencia de los aristotélicos y de los últimos científicos medievales, que se habían anclado en la consideración de cualidades que estimaban fundamentales, y estudiaban la adquisición y pérdida de éstas, o debatían el significado de las mismas, Galileo propuso buscar axiomas *cuantitativos*. Este cambio es muy importante; veremos todo su significado más adelante, pero puede ser útil ahora considerar un ejemplo elemental. Los aristotélicos decían que una bola cae porque tiene peso, que cae hacia la tierra porque todo objeto busca su sitio natural, y que para los cuerpos pesados el sitio natural es el centro de la tierra. Estos principios son cualitativos. Incluso la primera ley del movimeinto de Kepler, la de que la trayectoria de cada planeta es una elipse, es una afirmación cualitativa. En contraste, consideremos la afirmación de que la velocidad (en pies por segundo) con la que cae una bola es 32 veces el número de segundos que ha estado cayendo o, en símbolos, $v = 32t$. Esta es una afirmación cuantitativa acerca de cómo cae una bola. Galileo pretendía buscar sus axiomas como tales afirmaciones cuantitativas, y esperaba deducir algunos nuevos por medios matemáticos. Estas deducciones también proporcionarían un conocimiento cuantitativo. Además, como hemos visto, las matemáticas iban a ser su medio esencial.

La decisión de buscar un conocimiento cuantitativo expresado en fórmulas llevaba consigo otra decisión radical, aunque el primer contacto con ella difícilmente revela todo su significado. Los aristotélicos creían que una de las tareas de la ciencia era el explicar por qué sucedían las cosas; esta explicación significaba desvelar las causas del fenómeno. La explicación de que un cuerpo cae porque tiene peso proporciona la causa efectiva de la caída, y la afirmación de que busca su sitio natural expone la causa final. Pero el planteamiento cuantitativo $v = 32t$, independientemente de su valor, no explica por qué cae una bola, dice sólo cómo cambia la velocidad con el tiempo. En otras

palabras, las fórmulas no explican: describen. El conocimiento de la naturaleza que buscaba Galileo era descriptivo. Dice en *Dos nuevas ciencias:* «La causa de la aceleración del movimiento de los cuerpos que caen no es una parte necesaria de la investigación.» Más en general, dice que investigará y demostrará algunas de las propiedades del movimiento sin considerar cuáles pueden ser sus causas. La búsqueda científica positiva iba a separarse de las cuestiones de motivación última, e iba a abandonarse la especulación sobre las causas físicas.

Las primeras reacciones a este principio de Galileo posiblemente serían negativas. La descripción de los fenómenos en términos de fórmulas difícilmente podía parecer más que un primer paso. Parecería que la verdadera función de la ciencia, la de explicar por qué sucedían los fenómenos, había sido realmente comprendida por los aristotélicos. Incluso Descartes había protestado ante la decisión de Galileo de buscar fórmulas descriptivas. Dijo: «Todo lo que dice Galileo acerca de los cuerpos que caen en el espacio vacío está construido sin cimientos: debería haber determinado primero la naturaleza del peso.» Además, decía Descartes, Galileo debería reflexionar sobre las razones últimas. Pero veremos claramente después de algunos capítulos que la decisión de Galileo de buscar la descripción fue la idea más profunda y fructífera que se haya podido tener sobre el método científico.

Mientras que los aristotélicos habían hablado en términos de cualidades tales como fluidez, rigidez, esencias, lugares naturales, movimiento natural y violento y potencialidad, Galileo escogió un conjunto de conceptos enteramente nuevo, el cual, además, era medible, de modo que sus medidas podían relacionarse mediante fórmulas. Algunos de ellos son: distancia, tiempo, velocidad, aceleración, fuerza, masa y peso. Estos conceptos son demasiado familiares para sorprendernos. Pero en la época de Galileo eran elecciones radicales, al menos como conceptos fundamentales; y éstos fueron los que se revelaron como más operativos en la tarea de entender y dominar la naturaleza.

Hemos mostrado las características esenciales del programa de Galileo. Algunas de sus ideas ya las mantenían otros; algunas otras eran completamente originales suyas. Pero lo que establece la grandeza de Galileo es que vio con tanta claridad lo que estaba equivocado o era deficiente en los esfuerzos científicos de entonces, se liberó completamente de los antiguos modos y formuló tan claramente los

nuevos métodos. Además, al aplicarlos a los problemas del movimiento no sólo ejemplificó el método, sino que consiguió obtener resultados brillantes —en otras palabras, demostró que el método funcionaba—. La unidad de su trabajo, la claridad de sus pensamientos y expresiones y la fuerza de su argumentación influyó en casi todos sus contemporáneos y sucesores. Más que cualquier otro, Galileo es el fundador del método de la ciencia moderna. Era completamente consciente de lo que había realizado (ver el lema del capítulo), y también lo eran otros. El filósofo Hobbes dijo de Galileo: «Ha sido el primero en abrirnos la puerta del reino de la física.»

No podemos continuar la historia del método de la ciencia. Sin embargo, como las matemáticas se hicieron tan importantes en esta metodología y se aprovecharon tanto de su adopción, debemos hacer notar hasta qué punto el programa de Galileo fue aceptado completamente por gigantes como Newton. Este afirma que se necesitan los experimentos para obtener las leyes básicas. También es claro para él que la función de la ciencia, después de haber obtenido algunos principios básicos, es deducir nuevos hechos a partir de esos principios. En el prefacio de sus *Principia,* dice:

Puesto que los científicos (como nos dice Pappus) valoraban la ciencia de la mecánica como de la mayor importancia en la investigación de las cosas naturales, y los modernos, rechazando las formas sustanciales y las cualidades ocultas, se han esforzado por someter los fenómenos de la naturaleza a las leyes de las matemáticas, en este tratado he cultivado las matemáticas hasta donde se relacionan con la filosofía (ciencia)... y, en consecuencia, presento este trabajo como los principios matemáticos de la filosofía, porque la auténtica carga de la filosofía parece consistir en esto —a partir de los fenómenos del movimiento, investigar las fuerzas de la naturaleza, y entonces, mediante esas fuerzas, mostrar los otros fenómenos...

Por supuesto, principios matemáticos, para Newton y Galileo, eran principios cuantitativos. Dice en los *Principia* que su propósito es descubrir y establecer la forma exacta en la que «todas las cosas han sido ordenadas en medida, número y peso». Newton tenía una buena razón para hacer hincapié en las leyes matemáticas cuantitativas como contrapuestas a la explicación física, porque el concepto físico central en su mecánica celeste era la fuerza de la gravitación, cuya acción no podía explicarse en absoluto en términos físicos. En lugar de una explicación, Newton tenía una formulación cuantitativa de cómo actuaba la gravedad, que era significativa y utilizable. Y por ello dice,

al comienzo de los *Principia:* «Porque aquí pretendo sólo dar una noción matemática de estas fuerzas, sin considerar sus causas físicas y sedes.» Hacia el final del libro repite este pensamiento:

Pero nuestro propósito es sólo descubrir la cantidad y propiedades de esta fuerza a partir de los fenómenos, y aplicar lo que descubrimos en algunos casos sencillos, como principios, mediante los cuales, de forma matemática, podemos estimar los efectos de ellos en casos más complicados... Decimos, de *forma matemática* (itálica en Newton), para evitar todas las cuestiones sobre la naturaleza y cualidades de esta fuerza, que no pretenderíamos determinar mediante ninguna hipótesis...

El abandono del procedimiento físico en favor de la descripción matemática escandalizó incluso a grandes científicos. Huygens consideró la idea de la gravitación como «absurda», porque su acción a través del espacio vacío imposibilitaba cualquier acción mecánica. Mostró su sorpresa porque Newton se hubiera tomado el trabajo de realizar esa cantidad de cálculos laboriosos con el único fundamento del principio matemático de la gravitación. Leibniz atacó la gravitación como una potencia incorpórea e inexplicable; Jean Bernoulli (hermano de Jacques) la denunció como «repugnante a las mentes acostumbradas a no aceptar ningún principio en física que no fuera incontestable y evidente». Pero esta confianza en la descripción matemática, aun donde la comprensión física faltaba completamente, permitió a Newton sorprendentes contribuciones, por no mencionar los desarrollos siguientes.

Como la ciencia se hizo muy dependiente de las matemáticas, casi subordinada a ellas, fueron los científicos quienes extendieron el campo y las técnicas de las matemáticas, y la multiplicidad de problemas que suministró la ciencia proporcionó a los matemáticos muchas y profundas direcciones de trabajo creativo.

4. El concepto de función

El primer avance matemático que se obtuvo de las investigaciones científicas realizadas de acuerdo con el programa de Galileo provino del estudio del movimiento. Este problema absorbió a los científicos y matemáticos del siglo XVII. Es fácil ver por qué. Aunque la astronomía de Kepler fue aceptada ya a principios del siglo XVII, especialmen-

te después de que las observaciones de Galileo dieran pruebas adicionales de la teoría heliocéntrica, la ley de Kepler del movimiento elíptico es sólo aproximadamente correcta, si bien sería exacta si sólo estuvieran en el espacio el Sol y un planeta. Las ideas de que los otros planetas perturban el movimiento elíptico de cualquier otro planeta y de que el Sol perturba el movimiento elíptico de la Luna alrededor de la Tierra, ya habían sido consideradas; de hecho, la noción de una fuerza gravitacional actuando entre dos cuerpos cualesquiera fue sugerida por Kepler, entre otros. Por lo tanto, el problema de mejorar el cálculo de las posiciones de los planetas estaba abierto. Además, Kepler había obtenido sus leyes esencialmente ajustando curvas a los datos astronómicos, sin ninguna explicación, en términos de leyes fundamentales del movimiento, de por qué los planetas se movían en trayectorias elípticas. El problema básico de deducir las leyes de Kepler a partir de los principios del movimiento constituía un claro desafío.

La mejora de la teoría astronómica también tenía un objetivo práctico. En su búsqueda de materias primas y comercio, los europeos habían emprendido navegaciones en gran escala que implicaban el recorrido de grandes distancias fuera de la vista de tierra. Los marineros necesitaban, por tanto, métodos precisos para determinar la latitud y la longitud. La determinación de la latitud puede hacerse mediante observación directa del Sol o de las estrellas, pero la determinación de la longitud es bastante más difícil. En el siglo XVI, los métodos para hacerlo eran tan imprecisos que los navegantes cometían errores de hasta 500 millas. Después de, aproximadamente, 1514, se utilizaba la dirección de la Luna con relación a las estrellas para determinar la longitud. Estas direcciones, según se ven desde un mismo lugar de referencia en diferentes momentos, estaban tabuladas. Un navegante determinaría la dirección de la Luna, la cual no se vería muy afectada por estar en un lugar distinto, y determinaría su hora local utilizando, por ejemplo, las direcciones de las estrellas. Directamente de las tablas, o mediante interpolación, podía obtener la hora en el lugar de referencia cuando la Luna tenía la dirección medida y así calcular la diferencia de hora entre su posición y la de referencia. Cada hora de diferencia significa una diferencia de 15 grados en longitud. Este método, sin embargo, no era preciso. Como los barcos de aquellos tiempos estaban constantemente cabeceando, era difícil obtener en forma precisa la dirección de la Luna; pero como la Luna no se mueve mucho con relación a las estrellas en

pocas horas, la dirección de la Luna tenía que ser determinada con bastante precisión. Una equivocación de un minuto de ángulo significa un error de medio grado de longitud; pero incluso el realizar una medida con un error de menos de un minuto estaba lejos de las posibilidades de la época. Aunque fueron sugeridos e intentados otros métodos para determinar la longitud, parecía indispensable un mejor conocimiento de la trayectoria de la Luna para ampliar y mejorar las tablas, y muchos científicos, incluyendo Newton, trabajaron en el problema. Incluso en los tiempos de Newton, el conocimiento de la posición de la Luna era tan impreciso que el uso de las tablas conducía a errores de hasta 100 millas al determinar la posición en el mar.

Los gobiernos de Europa estaban muy interesados, porque las pérdidas en los envíos eran considerables. En 1675, el rey Carlos II de Inglaterra fundó el Royal Observatory en Greenwich para obtener mejores observaciones del movimiento de la Luna y para que sirviera como estación fija o de referencia para la determinación de la longitud. En 1672, el gobierno inglés estableció la Comission for the Discovery of Longitude, y ofreció recompensas de hasta 20.000 libras por ideas sobre cómo medir la longitud.

Los científicos del siglo XVII también se enfrentaron al problema de explicar los movimientos terrestres. Bajo la teoría heliocéntrica, la Tierra giraba sobre sí misma y efectuaba un movimiento de revolución en torno al Sol. ¿Por qué, entonces, deberían los objetos permanecer en ella? ¿Por qué los objetos que se tiran deberían caer hacia la Tierra, si ésta ya no era el centro del universo? Además, todos los movimientos, como el de los proyectiles, por ejemplo, parecían producirse como si la Tierra estuviera en reposo. Estas cuestiones atrajeron la atención de muchos, como Cardano, Tartaglia, Galileo y Newton. Las trayectorias de los proyectiles, sus alcances, la altura a que podían llegar, el efecto de la velocidad de la boca del arma sobre la altura y el alcance del proyectil, eran cuestiones básicas, y los príncipes entonces, como las naciones ahora, gastaron grandes sumas de dinero en sus soluciones. Se necesitaban nuevos principios del movimiento para tener en cuenta estos fenómenos terrestres, y los científicos pensaron que, como se creía que el universo estaba construido de acuerdo con un plan maestro, los mismos principios que explicaran los movimientos terrestres también lo harían con los celestes.

Del estudio de los varios problemas del movimiento emergió el más específico de diseñar métodos más precisos para medir el tiempo. Los relojes mecánicos, que se habían utilizado desde 1348, no eran

muy precisos. El cartógrafo flamenco Gemma Frisius había sugerido el uso de un reloj para determinar la longitud. Un barco podría llevar un reloj con la hora de un lugar de longitud conocida; como la determinación de la hora local, mediante la posición del Sol, por ejemplo, era relativamente sencilla, el navegante sólo necesitaba anotar la diferencia de hora y traducir ésta inmediatamente a diferencia de longitudes. Pero, incluso en 1600, no podía disponerse de relojes durables, precisos y en condiciones de navegar.

El movimiento de un péndulo parecía proporcionar el mecanismo básico para la medida del tiempo. Galileo había observado que el tiempo requerido para una oscilación completa de un reloj era constante, y ostensiblemente independiente de la amplitud de la oscilación. Preparó el diseño de un reloj de péndulo e hizo que su hijo construyera uno; pero fueron Robert Hooke y Huygens quienes realizaron el trabajo básico sobre el péndulo. Aunque el reloj de péndulo no era adecuado para un barco (se necesitaba una precisión de dos o tres segundos al día para el objetivo del cálculo de la longitud, y el movimiento del barco afectaba mucho a los péndulos), se reveló de valor inmenso en el trabajo científico, así como para la medida del tiempo en las casas y los negocios. Un reloj apropiado para la navegación fue finalmente diseñado por John Harrison (1693-1776) en 1761, y comenzó a utilizarse a fines del siglo XVIII. Como no fue posible disponer antes de un reloj propiamente dicho, la determinación precisa del movimiento de la Luna era todavía el principal problema científico en ese siglo.

Del estudio del movimiento obtuvieron las matemáticas un concepto fundamental, que fue central en prácticamente todo el trabajo de los siguientes doscientos años —el concepto de función o relación entre variables—. Se encuentra esta noción casi a lo largo de todo *Dos nuevas ciencias*, de Galileo, el libro en el que fundó la mecánica moderna. Galileo expresó sus relaciones funcionales en palabras y en el lenguaje de las proporciones. Así, en su trabajo sobre la resistencia de materiales, tiene ocasión de afirmar: «Las áreas de dos cilindros de volúmenes iguales, despreciando las bases, están una con respecto a otra en una razón que es la raíz cuadrada de la razón de sus longitudes.» En otra parte: «Los volúmenes de cilindros rectos cuyas superficies curvas son iguales son inversamente proporcionales a sus alturas.» En su trabajo sobre el movimiento establece, por ejemplo, que «los espacios descritos por un cuerpo que cae desde el reposo con un movimiento uniformemente acelerado están, unos con respecto a

otros, en la relación de los cuadrados de los intervalos de tiempo empleados en atravesar esas distancias». «Los tiempos de descenso a lo largo de planos inclinados de la misma altura, pero de diferentes pendientes, están unos con respecto a los otros en la relación de las longitudes de esos planos.» El lenguaje muestra claramente que está tratando con variables y funciones; no faltaba más que un paso para escribir estas frases en forma simbólica. Como el simbolismo del álgebra se estaba extendiendo en ese momento la afirmación de Galileo sobre los espacios descritos por un cuerpo que cae pronto se escribió como $s = kt^2$ y su afirmación sobre los tiempos de descenso como $t = kl$.

Muchas de las funciones introducidas durante el siglo XVII fueron estudiadas en primer lugar como curvas, antes de que el concepto de función fuera totalmente identificado. Esto ocurrió, por ejemplo, en el caso de las funciones trascendentes elementales tales como $\log x$, $\operatorname{sen} x$ y a^x. Así, Evangelista Torricelli (1608-1647), discípulo de Galileo, en una carta de 1644, describía sus investigaciones sobre la curva que nosotros representaríamos mediante $y = ae^{-cx}$ para $x \geqslant 0$ (el manuscrito en el que escribió esta investigación no fue editado hasta 1900). La curva le fue sugerida a Torricelli por el trabajo que se desarrollaba entonces sobre los logaritmos. Descartes encontró la misma curva en 1639, pero él no mencionó su conexión con los logaritmos. La curva seno apareció en las matemáticas como la curva asociada a la cicloide, en el trabajo de Roberval sobre la cicloide (cap. 17, sec. 2) y aparece dibujada a lo largo de dos períodos en la *Mechanica* de Wallis (1670). Las tablas de valores de las funciones trigonométricas y logarítmicas eran conocidas, por supuesto, en aquella época, con gran precisión.

Es también relevante el que las antiguas y las nuevas curvas fueran introducidas mediante movimientos. En la época griega, pocas curvas, como la cuadratriz y la espiral de Arquímedes, estaban definidas en términos de movimiento, pero en aquellos tiempos tales curvas estaban fuera de los límites de las matemáticas legítimas. La actitud era completamente diferente en el siglo XVII. Mersenne, en 1615, definió la cicloide (que era conocida anteriormente) como el lugar geométrico (que describe) un punto (fijo) de una rueda que gira sobre el suelo. Galileo, que había demostrado que la trayectoria de un proyectil disparado en el aire formando un ángulo con respecto al suelo es una parábola, consideró la curva como el lugar geométrico de un punto móvil.

Con Roberval, Barrow y Newton, el concepto de curva como la trayectoria de un punto móvil alcanza reconocimiento explícito y aceptación. Dice Newton en su *Quadrature of Curves* (escrito en 1676): «Considero las cantidades matemáticas en este punto no como constituidas por muy pequeñas partes, sino como descritas por un movimiento continuado. Las líneas [curvas] están descritas, y así generadas, no por la yuxtaposición de partes sino por el movimiento continuado de puntos... Esta génesis tiene lugar realmente en la naturaleza de las cosas, y se ve diariamente en el movimiento de los cuerpos.»

Los términos y el simbolismo para los distintos tipos de funciones representadas por estas curvas fueron introduciéndose gradualmente. Había muchas dificultades sutiles de las que casi no se era consciente. Por ejemplo, el uso de las funciones de la forma a^x, cuando x toma valores positivos y negativos, enteros y fraccionarios, llegó a ser común en el siglo XVII. Se suponía (hasta el siglo XIX, cuando se definieron por primera vez los números irracionales) que la función estaba también definida para valores irracionales de x, de modo que nadie cuestionaba una expresión de la forma $2^{\sqrt{2}}$. Se entendía implícitamente que un valor así era intermedio entre los obtenidos para dos exponentes racionales cualesquiera por encima y por debajo de $\sqrt{2}$.

La distinción de Descartes entre curvas geométricas y mecánicas (cap. 15, sec. 4) suscitó la distinción entre funciones algebraicas y trascendentes. Afortunadamente, sus contemporáneos ignoraron su rechazo de lo que él llamó curvas mecánicas. Mediante cuadraturas, sumación de series y otras operaciones incluidas en el cálculo, surgieron y fueron estudiadas muchas funciones trascendentes. La distinción entre funciones algebraicas y trascendentes fue hecha por James Gregory en 1667, cuando intentó demostrar que el área de un sector circular no podía ser una función algebraica del radio y de la cuerda. Leibniz demostró que sen x no podía ser una función algebraica de x e incidentalmente demostró el resultado buscado por Gregory [2]. La comprensión completa y el uso de las funciones trascendentes vino gradualmente.

La definición más explícita del concepto de función en el siglo XVII fue dada por James Gregory en su *Vera Circuli et Hyperbolae Quadratura* (1667). Definió una función como una cantidad

[2] *Math. Schriften*, 5, 97-98.

que se obtiene de otras cantidades mediante una sucesión de operaciones algebraicas o mediante cualquier otra operación imaginable. Con la última frase quería decir, según explica, que es necesario añadir a las cinco operaciones del álgebra una sexta operación, que él define como el paso al límite. (A Gregory, como veremos en el capítulo 17, le preocupaban los problemas de cuadraturas.) El concepto de función de Gregory no se conservó; pero, en cualquier caso, pronto se habría vuelto demasiado restringido al utilizarse, cada vez más ampliamente, la representación de funciones mediante series.

Desde el mismo comienzo de su trabajo sobre el cálculo, es decir, desde 1665 en adelante, Newton utilizó el término «fluent» (fluyente) para representar cualquier relación entre variables. En un manuscrito de 1673, Leibniz utilizó la palabra «función» para significar cualquier cantidad que varía de un punto a otro de una curva —por ejemplo, la longitud de la tangente, de la normal, de la subtangente y de la ordenada—. La curva misma se decía dada mediante una ecuación. Leibniz también introdujo las palabras «constante», «variable» y «parámetro», esta última utilizada en conexión con una familia de curvas [3]. Tratando con funciones, Jean Bernoulli hablaba ya desde 1697 de una cantidad formada, de cualquier manera posible, de variables y constantes [4]; con «cualquier manera» quería decir mediante expresiones algebraicas y trascendentes. Adoptó la frase de Leibniz «función de x» para esta cantidad en 1698. En su *Historia* (1714), Leibniz utilizó la palabra «función» para significar cantidades que dependen de una variable.

En cuanto a la notación, Jean Bernoulli escribía X ó ξ para una función general de x, aunque en 1718 cambió a ϕx. A Leibniz le pareció bien esto, pero propuso también x^1 y x^2 para funciones de x, utilizando el superíndice cuando se tratara con varias funciones. La notación $f(x)$ fue introducida por Euler en 1734 [5]. El concepto se convirtió inmediatamente en central en los trabajos sobre el cálculo. Veremos más adelante cómo fue extendido ese concepto.

---

[3] *Math. Schriften*, 5, 266-269.

[4] *Mém. de l'Acad. des Sci.*, París, 1718, 100 ss. = *Opera*, 2, 235-269, p. 241, en particular.

[5] *Comm. Acad. Sci. Petrop.*, 7, 1734/5, 184-200, pub. 1740 = *Opera* (1), 22, 57-75.

Bibliografía

Bell, A. E.: *Christian Huygens and the Development of Science in the Seventeenth Century*, Edward Arnold, 1947.

Burtt, E. A.: *Los fundamentos metafísicos de la Ciencia Moderna*, Editorial Suramericana, Buenos Aires, 1945.

Butterfield, Herbert: *The Origins of Modern Science*, Macmillan, 1951, caps. 4-7. H. Butterfield, *Los orígenes de la ciencia moderna*, Madrid, Taurus, 1982.

Cohen, I. Bernard: *The Birth of a New Physics*, Doubleday, 1960.

Coolidge, Julian L.: *The Mathematics of Great Amateurs*, Dover (reimpresión), 1963, pp. 119-127.

Crombie, A. C.: *Augustine to Galileo*, Falcon Press, 1952, Cap. 6. A. C. Crombie, *Historia de la ciencia: De San Agustín a Galileo*, Madrid, Alianza Editorial, 1979.

Dampier-Whetham, W. C. D.: *A History of Science and Its Relations with Philosophy and Religion*, Cambridge University Press, 1929, cap. 3. W. C. Dampier, *Historia de la ciencia y sus relaciones con la filosofía y religión*, Madrid, Tecnos, 1986.

Dijksterhuis, E. J.: *The Mechanization of the World Picture*, Oxford University Press, 1961.

Drabkin, I. E., and Stillman Drake: *Galileo Galilei: On Motion and Mechanics*, University of Wisconsin Press, 1960.

Drake, Stillman: *Discoveries and Opinions of Galileo*, Doubleday, 1957.

Galilei, Galileo: *Opere*, 20 vols., 1890-1909, reimpreso por G. Barbera, 1964-1966.

—: *Dialogues Concerning Two New Sciences*, Dover (reimpresión), 1952. *Consideraciones y demostraciones matemáticas sobre las nuevas ciencias*, Editora Nacional, Madrid, 1976.

Hall, A. R.: *The Scientific Revolution*, Longmans Green, 1954, caps. 1-8.

—: *From Galileo to Newton*, Collins, 1963, caps. 1-5. A. R. Hall, *La Revolución Científica*, 1500-1700, Barcelona, Crítica, 1985.

Huygens, C.: *Oeuvres complètes*, 22 vols., M. Nyhoff, 1888-1950.

Newton, I.: *Mathematical Principles of Natural Philosphy*, University of California Press, 1946. I. Newton, *Principios matemáticos de la filosofía material*, 2 vol., Madrid, Alianza, 1987.

Randall, John H., Jr.: *Making of the Modern Mind*, ed. rev., Houghton Mifflin, 1940, cap. 10.

Scott, J. F.: *The Scientific Work of René Descartes*, Taylor and Francis, 1952, caps. 10-12.

Smith, Preserved: *A History of Modern Culture*, Henry Holt, 1930, vol. I, caps. 3, 6 y 7.

Strong, Edward, W.: *Procedures and Metaphysics*, University of California Press, 1936; caps. 5-8.

Whitehead, Alfred North: *Science and Modern World*, Cambridge University Press, 1926, cap. 3.
Wolf, Abraham: *A History of Science, Technology and Philosophy in the Sixteenth and Seventeenth Centuries*, 2.ª ed., George Allen and Unwin, 1950, cap. 3.

Capítulo 17
LA CREACION DEL CALCULO

> Quien, por un vigor de la mente casi divino, los movimientos y las figuras de los planetas, las trayectorias de los cometas y las mareas de los mares primero demostró.
>
> EPITAFIO DE NEWTON

1. La motivación del cálculo

Justamente después de la adopción del concepto de función vino el cálculo, el cual, junto con la geometría euclídea, es la mayor creación de todas las matemáticas. Aunque era, hasta cierto punto, la respuesta a problemas ya manejados por los griegos, el cálculo fue creado sobre todo para tratar los principales problemas científicos del siglo XVII.

Había cuatro tipos principales de problemas. El primero era el siguiente: dada la fórmula de la distancia que un cuerpo recorre como función del tiempo, obtener la velocidad y la aceleración en cualquier instante; y, al revés, dada la fórmula que describe la aceleración de un cuerpo como función del tiempo, obtener la velocidad y la distancia recorrida. Este problema surgió directamente en el estudio del movimiento, y la dificultad que planteaba era que las velocidades y las aceleraciones que interesaban en el siglo XVII variaban de instante en instante. Al calcular una velocidad instantánea, por ejemplo, no se puede dividir la distancia recorrida por el tiempo empleado, como ocurre en el caso del cálculo de la velocidad media, porque en un instante dado tanto la distancia recorrida como el

tiempo empleado son cero, y 0/0 no tiene sentido. Sin embargo, era claro desde un punto de vista físico que los objetos móviles tienen una velocidad en cada instante de su viaje. El problema inverso de obtener la distancia recorrida, conociendo la fórmula de la velocidad, incluye la dificultad correspondiente; no se puede multiplicar la velocidad en cualquier instante por el tiempo utilizado para obtener el espacio recorrido, porque la velocidad varía de un instante a otro.

El segundo tipo de problemas era obtener la tangente a una curva. El interés por este problema vino de más de una fuente; era un problema de geometría pura, y era de gran importancia para las aplicaciones científicas. La óptica, como sabemos, era uno de los principales objetivos científicos del siglo XVII; el diseño de las lentes era de interés directo para Fermat, Descartes, Huygens y Newton. Para estudiar el paso de la luz a través de una lente, se debe conocer el ángulo bajo el cual el rayo toca a la lente, para aplicar la ley de refracción. El ángulo significativo es el que forman el rayo y la normal a la curva (fig. 17.1), donde la normal es la perpendicular a la tangente. Por lo tanto el problema era obtener o la normal o la tangente. Otro problema científico que implicaba la tangente a una curva surgía en el estudio del movimiento. La dirección del movimiento de un cuerpo móvil en cualquier punto de su trayectoria es la dirección de la tangente a la trayectoria.

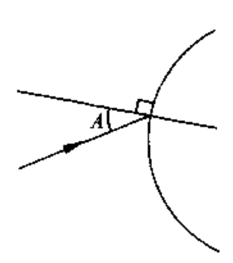


FIGURA 17.1

En realidad, incluso el mismo significado de «tangente» estaba abierto. Para las secciones cónicas, la definición de una tangente como una recta que toca a una curva en sólo un punto y que permanece a un lado de la curva, bastaba; esta definición era utilizada por los griegos. Pero era inadecuada para las curvas más complicadas que se utilizaban en el siglo XVII.

El tercer problema era obtener el valor máximo o mínimo de una

función. Cuando una bala se dispara desde un cañon, la distancia que recorrerá horizontalmente —el recorrido— depende del ángulo de inclinación del cañón con respecto al suelo. Un problema «práctico» era obtener el ángulo que haría máximo el recorrido. A principios del siglo XVII Galileo obtuvo que (en el vacío) el recorrido máximo se obtenía para un ángulo de fuego de 45°; también encontró la máxima altura alcanzada por proyectiles disparados con distintos ángulos respecto al suelo. El estudio del movimiento de los planetas también presentaba problemas de máximos y mínimos, tales como los de obtener la mayor y la menor distancia de un planeta al Sol.

El cuarto problema era el de obtener longitudes de curvas como, por ejemplo, la distancia recorrida por un planeta en un período de tiempo dado; las áreas acotadas por curvas; los volúmenes acotados por superficies; los centros de gravedad de los cuerpos y la atracción gravitatoria que un cuerpo *extenso*, un planeta, por ejemplo, ejerce sobre otro cuerpo. Los griegos habían aplicado el método exhaustivo para obtener algunas áreas y volúmenes. A pesar del hecho de que lo aplicaban para áreas y volúmenes relativamente sencillos, tenían que utilizar mucha ingeniosidad, porque al método le faltaba generalidad, y no obtuvieron respuestas numéricas muy a menudo. El interés por obtener longitudes, áreas, volúmenes y centros de gravedad revivió cuando los trabajos de Arquímedes se hicieron conocidos en Europa. El método exhaustivo se modificó primero gradualmente, y después radicalmente por la invención del cálculo.

2. El trabajo sobre el cálculo de principios del siglo XVII

Los problemas del cálculo fueron abordados por, al menos, una docena de los matemáticos más grandes del siglo XVII, y por varias docenas de otros menos importantes. Todas sus contribuciones fueron coronadas por las realizaciones de Newton y Leibniz. Aquí podremos señalar sólo las contribuciones principales de los precursores de estos dos maestros.

El problema del cálculo de la velocidad instantánea a partir del conocimiento de la distancia recorrida como función del tiempo, y su inverso, se vio pronto que eran casos particulares del cálculo del cambio relativo instantáneo de una variable con respecto a otra, y su problema inverso. El primer tratamiento significativo de los pro-

blemas de cambios relativos, en general, se debe a Newton; lo examinaremos más tarde.

Se propusieron varios métodos para obtener la tangente a una curva. En su *Traité des indivisibles,* que data de 1634 (aunque no fue publicado hasta 1693), Giles Persone de Roberval (1602-1675) generalizó un método que Arquímedes había usado para obtener la tangente en cualquier punto de su espiral. Como Arquímedes, Roberval pensó en una curva como lugar geométrico de un punto que se mueve bajo la acción de dos velocidades. Así, un proyectil disparado desde un cañón experimenta la acción de una velocidad horizontal, *PQ* en la figura 17.2, y una velocidad vertical, *PR*. La resultante de estas dos velocidades es la diagonal del rectángulo formado sobre *PQ* y *PR*. Roberval tomó la recta de esta diagonal como la tangente en *P*. Como señaló Torricelli, el método de Roberval utilizaba un principio establecido ya por Galileo, que consiste en que las velocidades horizontal y vertical actúan independientemente la una de la otra. El mismo Torricelli utilizó el método de Roberval para obtener las tangentes a las curvas cuyas ecuaciones escribimos en la actualidad como $y = x^n$.

Aunque la noción de tangente como una recta que tiene la dirección de la velocidad resultante era más complicada que la definición griega de una recta que toca a una curva, el nuevo concepto podía aplicarse a muchas curvas en el que el antiguo fallaba. Era también valioso porque relacionaba la geometría pura y la dinámica, las cuales, antes de los trabajos de Galileo, habían sido consideradas como esencialmente distintas. Por otra parte, esta definición de tangente era objetable en términos matemáticos, porque se basaba en conceptos físicos. Surgieron muchos casos de curvas en situaciones que no se podían relacionar con el movimiento y por lo tanto la definición de tangente no se podía aplicar. Por todo ello fueron ganando aceptación otros métodos para obtener tangentes.

El método de Fermat, que él había ideado en 1629 y que puede encontrarse en su manuscrito de 1637 *Methodus ad Disquirendam Maximam et Minimam* (Métodos para obtener máximos y mínimos) [1], es, en esencia, el método actual. Sea *PT* la tangente deseada a una curva en *P* (fig. 17.3). La longitud *TQ* se llama subtangente. El plan de Fermat es obtener la longitud de *TQ*, de la que se obtiene la posición de *T* y entonces trazar *TP*.

[1] *Oeuvres,* 1, 133-179; 3, 121-156.

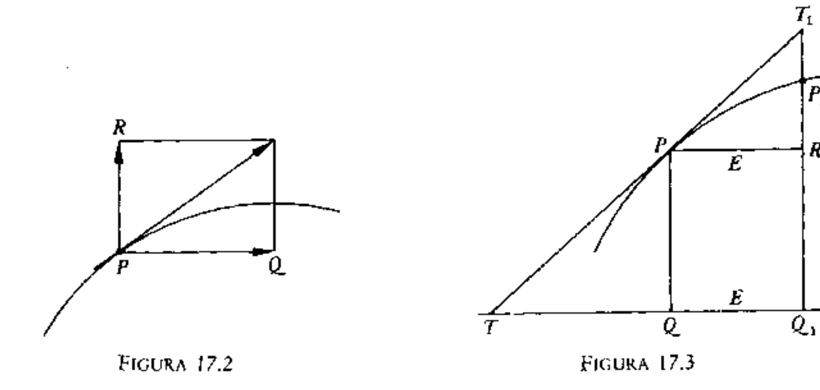


FIGURA 17.2

FIGURA 17.3

Sea QQ_1 un incremento de longitud E de TQ. Como el triángulo TQP es semejante al triángulo PRT_1,

$$TQ : PQ = E : T_1R.$$

Pero, dice Fermat, T_1R es casi P_1R, por lo tanto

$$TQ : PQ = E : (P_1Q_1 - QP).$$

Llamando $f(x)$ a PQ, en nuestra notación moderna, tenemos

$$TQ : f(x) = E : [f(x + E) - f(x)].$$

Por tanto

$$TQ = \frac{Ef(x)}{f(x + E) - f(x)}.$$

Para la $f(x)$ tratada por Fermat era inmediatamente posible dividir el numerador y el denominador de la fracción anterior por E. Hace entonces $E/ = 0$ (según dice, elimina el término E) y así obtiene TQ.

Fermat aplicó este método de las tangentes a muchos problemas difíciles. El método tiene la *forma* del método, ahora habitual, del cálculo diferencial, pero está suponiendo enteramente la difícil teoría de los límites.

Para Descartes, encontrar la tangente a una curva era importante porque permitía encontrar propiedades de las curvas —como, por ejemplo, el ángulo de intersección de dos de ellas—. Dice: «este es el problema más útil, y el más general, no sólo que conozco, sino de los que deseo conocer en geometría». Escribió su método en el segundo libro de *La Géométrie.* Era puramente algebraico y no incluía ningún concepto de límite, mientras que para Fermat sí lo implicaba, si se formulaba rigurosamente. Sin embargo, el método de Descartes sólo era útil para curvas cuyas ecuaciones fueran de la forma $y = f(x)$, donde $f(x)$ era un polinomio sencillo. Aunque el método de Fermat era general, Descartes pensaba que su método era mejor; criticó el de Fermat, del que hay que reconocer que no resultaba claro en su presentación de entonces, e intentó interpretarlo en términos de sus propias ideas. Fermat, a su vez, proclamaba que su método era superior y veía ventajas en su utilización de los pequeños incrementos E.

Isaac Barrow (1630-1677) también dio un método para obtener tangentes a las curvas. Barrow era un profesor de matemáticas de la universidad de Cambridge. Muy versado en griego y árabe, pudo traducir algunos de los trabajos de Euclides y mejorar algunas otras traducciones de los escritos de Euclides, Apolonio, Arquímedes y Teodosio. Su trabajo más importante, las *Lectiones Geometricae* (1669), es una de las grandes contribuciones al cálculo. En él utilizaba métodos geométricos, «liberados», según decía, «de las abominables cargas del cálculo». En 1669, Barrow renunció a su puesto de profesor en favor de Newton y volvió a los estudios teológicos.

El método geométrico de Barrow es bastante complicado y hace uso de curvas auxiliares. Sin embargo, vale la pena destacar una característica, porque ilustra la forma de pensar de la época; y es el uso de lo que se llama el triángulo diferencial, o característico. Comienza considerando el triángulo PRQ (fig. 17.4), que se obtiene por el incremento PR, y utiliza el hecho de que este triángulo es semejante al PMN para afirmar que la pendiente QR/PR de la tangente es igual a PM/MN. Sin embargo, dice Barrow, cuando el arco PP' es suficientemente pequeño podemos identificarlo sin gran error con el segmento PQ de la tangente en P. El triángulo PRP' (fig. 17.5), en el que PP' puede considerarse bien como un arco de la curva o como una parte de la tangente, es el triángulo característico. Había sido utilizado mucho antes por Pascal, en conexión con la obtención de áreas, y por otros antes que él.

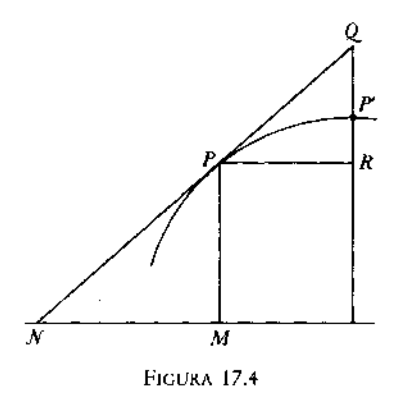


FIGURA 17.4

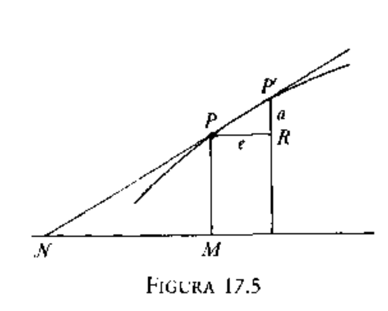


FIGURA 17.5

En la lección 10 de las *Lectiones*, Barrow recurre al cálculo para obtener la tangente a una curva. Aquí el método es esencialmente el mismo que el de Fermat. Utiliza la ecuación de la curva, por ejemplo $y^2 = px$, y sustituye x por $x + e$ e y por $y + a$. Entonces

$$y^2 + 2ay + a^2 = px + pe.$$

Resta $y^2 = px$ y obtiene

$$2ay + a^2 = pe.$$

A continuación desprecia las potencias superiores de a y e (cuando aparecen), que es lo mismo que sustituir PRP' de la figura 17.4 por PRP' de la figura 17.5, y concluye que

$$\frac{a}{e} = \frac{p}{2y}.$$

Ahora bien, como $a/e = PM/NM$, se tiene

$$\frac{PM}{NM} = \frac{p}{2y}.$$

Como PM es y, ha calculado la subtangente NM, y conoce entonces la posición de N.

El trabajo sobre el tercer tipo de problemas, la obtención de los máximos y mínimos de una función, puede decirse que comienza

con una observación de Kepler. Estaba interesado en la forma de los toneles de vino; en su *Stereometria Doliorum* (1615) demostró que, de todos los paralelepípedos rectos, de bases cuadradas, inscritos en una esfera, el cubo es el mayor. Su método fue el de calcular los volúmenes para elecciones particulares de las dimensiones. Esto, en sí mismo, no era significativo; pero notó que cuando se acercaba al volumen máximo, el *cambio* en volumen que correspondía a un cambio fijo en las dimensiones crecía cada vez menos.

Fermat, en su *Methodus ad Disquirendam*, describió su método, que ilustró mediante el siguiente ejemplo: dado un segmento, se desea encontrar un punto de él tal que el rectángulo formado con los dos segmentos en que queda dividido sea máximo. Llama B a todo el segmento, y sea A una parte de él. El área del rectángulo formado por los dos segmentos es $AB - A^2$. Ahora sustituye A por $A + E$. La otra parte es, entonces, $B - (A + E)$, y el área del rectángulo es ahora $(A + E)(B - A - E)$. Iguala las dos áreas porque, según razona, en un máximo los dos valores de la función —es decir, las dos áreas— deben de ser iguales. Por tanto,

$$AB + EB - A^2 - 2AE - E^2 = AB - A^2.$$

Restando los términos comunes en los dos miembros y dividiendo por E, obtiene

$$B = 2A + E.$$

Hace entonces $E = 0$ (él dice que desprecia el término en E) y obtiene $B = 2A$. Por lo tanto, el rectángulo es un cuadrado.

El método, en palabras de Fermat, es bastante general; él lo describe así: si A es la variable independiente, y si A se incrementa hasta $A + E$, entonces cuando E se hace indefinidamente pequeño y cuando la función pasa por un máximo o un mínimo, los dos valores de la función han de ser iguales. Estos valores se igualan; la ecuación se divide por E, y E se hace tender a cero, de modo que puede determinarse a partir de la ecuación el valor de A que hace máxima o mínima a la función. El método es esencialmente el que utilizaba para obtener la tangente a una curva. Sin embargo, el hecho básico allí es la semejanza de dos triángulos; aquí es la igualdad de dos valores de la función. Fermat no vio la necesidad de justificar la

introducción de un E distinto de cero y después, después de dividir por E, hace $E = 0$ [2].

Los trabajos del siglo XVII sobre obtención de áreas, volúmenes, centros de gravedad y longitudes de curvas comienzan con Kepler, de quien se dice que se interesó por el problema de los volúmenes porque notó la falta de precisión de los métodos utilizados por los tratantes en vinos para obtener el volumen de los barriles. Este trabajo (en *Stereometria Doliorum*) es tosco para los niveles actuales. Por ejemplo, el área de un círculo es, para él, el área de un número infinito de triángulos, cada uno con un vértice en el centro y una base en la circunferencia. De la fórmula del área de un polígono regular inscrito en una circunferencia, la mitad del perímetro por el apotema, obtenía el área del círculo. De forma análoga, consideraba el volumen de una esfera como la suma de los volúmenes de pequeños conos cuyos vértices están en el centro de la esfera y cuyas bases están en la superficie de la esfera. Así demostró que el volumen de la esfera es un tercio del radio por la superficie. Consideró el cono como una suma de discos circulares muy estrechos y pudo así calcular su volumen. Estimulado por la obra de Arquímedes *Esferoides y Conoides*, generó nuevas figuras mediante rotación de áreas y calculó los volúmenes correspondientes. Así calculó el volumen de la figura generada por el giro alrededor de su cuerda de un segmento circular.

La identificación de las áreas y volúmenes curvilíneos con la suma de un número infinito de elementos infinitesimales es la esencia del método de Kepler. El que el círculo pudiera considerarse como la suma de un número infinito de triángulos estaba justificado, para él, por el principio de continuidad (cap. 14, sec. 5). No veía ninguna diferencia de principio entre las dos figuras. Por la misma razón, una línea y un área infinitesimal eran realmente lo mismo; y de hecho consideró, en algunos problemas, un área como suma de líneas.

En *Dos nuevas ciencias* Galileo concibe las áreas en una forma parecida a Kepler; al tratar el problema del movimiento uniformemente acelerado, presentó un razonamiento para mostrar que el área encerrada bajo la curva tiempo-velocidad es la distancia. Suponga-

[2] Para las ecuaciones que preceden al hacer $E = 0$, Fermat utilizó el término *adaequalitas*, que Carl. B. Boyer en *The Concepts of the Calculus*, p. 156, ha traducido adecuadamente como «pseudo-igualdad».

mos un objeto que se mueve con velocidad variable $v = 32t$, representado por la línea recta de la figura 17.6; entonces, la distancia recorrida en el tiempo OA es el área OAB. Galileo llegó a esta conclusión considerando, por ejemplo, $A'B'$ como una velocidad típica en un instante y también como la distancia infinitesimal recorrida (como sería si se multiplicara por un elemento de tiempo muy pequeño), y razonando entonces que el área OAB, que está construida con las líneas $A'B'$, debe de ser, por tanto, la distancia total. Como AB es $32t$ y OA es t, el área OAB es $16t^2$. El razonamiento es, por supuesto, poco claro. Estaba apoyado en la mente de Galileo por consideraciones filosóficas que equivalían a considerar el área OAB como construida con un número infinito de unidades indivisibles como $A'B'$. Dedicó mucho tiempo al problema de la estructura de magnitudes continuas como los segmentos de rectas y las áreas, pero no lo resolvió.

Bonaventura Cavalieri (1598-1647), discípulo de Galileo y profesor en un liceo de Bolonia, fue influido por Kepler y Galileo y fue estimulado por este último para interesarse por problemas del cálculo. Cavalieri desarrolló las ideas de Galileo y otros sobre los indivisibles mediante un método geométrico, y publicó un trabajo sobre el tema, *Geometria Indivisibilibus Continuorum Nova quadam Ratione Promota* (Geometría superior mediante un método bastante desconocido, los indivisibles de los continuos, 1635). Considera un área como constituida por un número indefinido de rectas paralelas y equidistantes y un volumen como compuesto por un número indefinido de áreas planas paralelas; a estos elementos los llama los indivisibles de área y volumen, respectivamente. Cavalieri es consciente de que el número de indivisibles que constituyen un área o un volumen debe ser infinitamente grande, pero no trata de profundizar en esto. En líneas generales, los indivisibilistas mantenían, como expresara Cavalieri en sus *Exercitationes Geometricae Sex* (1647), que una línea está hecha de puntos como una sarta de cuentas; un plano está hecho de líneas, como un tejido de hebras, y un sólido de áreas planas como un libro de hojas. Sin embargo, aceptaban un número infinito de elementos constituyentes.

El método o principio de Cavalieri puede ilustrarse mediante la proposición siguiente que, por supuesto, puede demostrarse de otras formas. Para demostrar que el paralelogramo $ABCD$ (fig. 17.7) tiene área doble que cualquiera de los triángulos ABD o BCD, hace notar que cuando $GD = BE$, se tiene que $GH = FE$. Por tanto, los trián-

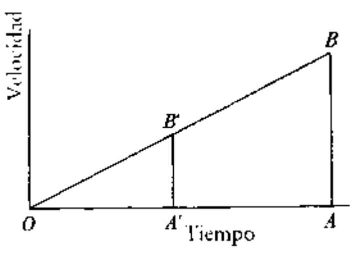


FIGURA 17.6

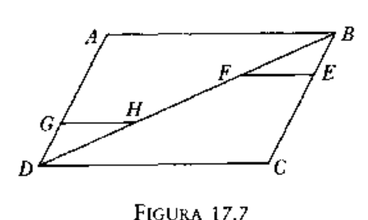


FIGURA 17.7

gulos *ABD* y *BCD* están constituidos por igual número de líneas iguales, tales como *GH* y *EF*, y por tanto tienen que tener áreas iguales.

Este mismo principio está incluido en la proposición que se enseña actualmente en los libros de geometría de sólidos y que se conoce como teorema de Cavalieri. El principio establece que si dos sólidos tienen igual altura y si las secciones por planos paralelos a las bases y a la misma distancia de ellas siempre están en una razón dada, los volúmenes de los dos sólidos también están en esa razón. Utilizando esencialmente este principio, Cavalieri demostró que el volumen de un cono es 1/3 del volumen del cilindro circunscrito. Trató de la misma forma el área limitada bajo dos curvas, $y = f(x)$ e $y = g(x)$, en nuestra notación, y definidas para los mismos valores de x; considerando las áreas como la suma de las ordenadas, si las ordenadas de una están en una razón constante con respecto a las de la otra entonces, según Cavalieri, las áreas están en la misma razón. Demostró mediante sus métodos en *Centuria di varii problemi* (1639) que, en nuestra notación,

$$\int_0^a x^n dx = \frac{a^{n+1}}{n+1}$$

para valores enteros positivos de n hasta 9. Sin embargo, su método era enteramente geométrico. Logró obtener resultados correctos porque aplicó su principio para calcular áreas y volúmenes en los que la razón de los indivisibles que constituían las respectivas áreas y volúmenes era constante.

La teoría de los indivisibles de Cavalieri fue criticada por sus contemporáneos, y Cavalieri intentó responderles, pero no tenía ninguna justificación rigurosa. A veces pretendía que su método era

sólo un instrumento pragmático para evitar el método exhaustivo. A pesar de las críticas al método, éste fue utilizado intensamente por muchos matemáticos. Otros, como Fermat, Pascal y Roberval, utilizaron el método e incluso la nomenclatura, como la suma de ordenadas, pero consideraban el área como una suma de infinitos rectángulos pequeños en lugar de una suma de líneas.

En 1634, Roberval, quien dice que había estudiado al «divino Arquímedes», utilizó esencialmente el método de los indivisibles para obtener el área encerrada bajo un arco de cicloide, un problema sobre el que Mersenne había llamado su atención en 1629. A Roberval se le acredita a veces el descubrimiento independiente del método de los indivisibles, pero en realidad él creía en la infinita divisibilidad de las líneas, superficies y volúmenes, de manera que no habría partes últimas. Llamó a su método el «método de las infinidades», aunque utilizó como título de su trabajo el de *Traité des indivisibles*.

El método de Roberval para obtener el área encerrada por la cicloide es instructivo. Sea $OABP$ (fig. 17.8) el área situada bajo la mitad de un arco de cicloide. El diámetro de la circunferencia generatriz es OC y P es un punto cualquiera del arco. Se toma $PQ = DF$. El lugar geométrico descrito por Q se llama curva asociada a la cicloide. (La curva OQB es, en nuestra notación, $y = a$ sen (x/a), donde a es el radio de la circunferencia generatriz, con tal que el origen esté en el punto medio de OQB y el eje OX sea paralelo a OA.) Roberval afirma que la curva OQB divide al rectángulo $OABC$ en dos partes iguales porque, básicamente, a cada línea DQ en $OQBC$ le corresponde una línea igual RS en $OABQ$. Entonces puede aplicarse el principio de Cavalieri. El rectángulo $OABC$ tiene su base y altura iguales, respectivamente, a la semicircunferencia y diámetro de la circunferencia generatriz; por lo tanto su área es doble de la de la circunferencia. Entonces $OABQ$ tiene la misma área que la circunferencia generatriz. Además, el área entre OPB y OQB es igual al área del semicírculo OFC porque de la misma definición de Q se tiene que $DF = PQ$, de modo que estas dos áreas tienen la misma anchura en todas partes. En consecuencia, el área encerrada debajo del semiarco es una vez y media el área de la circunferencia generatriz. Roberval también obtuvo el área encerrada en un arco de la curva seno, el volumen generado por la revolución del arco alrededor de su base, otros volúmenes conectados con la cicloide y el centroide de su área.

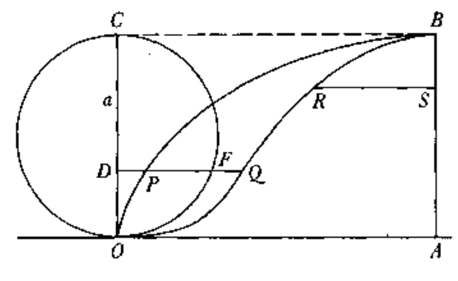


Figura 17.8

El método nuevo más importante para calcular áreas, volúmenes y otras cantidades comenzó con modificaciones del método exhaustivo griego. Consideremos un ejemplo típico. Supongamos que se desea calcular el área situada debajo de la parábola $y = x^2$ desde $x = 0$ hasta $x = B$ (fig. 17.9). Mientras que el método exhaustivo utilizaba diferentes tipos de figuras aproximantes rectilíneas, dependiendo del área curvilínea en cuestión, algunos adoptaron en el siglo XVII un procedimiento sistemático utilizando rectángulos, como se muestra en la figura. A medida que la anchura d de estos rectángulos se hace más pequeña, la suma de las áreas de los rectángulos se aproxima al área encerrada bajo la curva. Esta suma, si las bases son todas ellas de anchura d, y si se utiliza la propiedad característica de la parábola de que la ordenada es el cuadrado de la abscisa, es

$$d \cdot d^2 + d(2d)^2 + d(3d)^2 + \dots + d(nd)^2 \qquad (1)$$

o

$$d^3(1 + 2^2 + 3^2 + \dots + n^2).$$

Ahora bien, la suma de las potencias *m-ésimas* de los primeros n números naturales había sido obtenida por Pascal y Fermat precisamente para su uso en tales problemas; por ello los matemáticos pudieron sustituir fácilmente la última expresión por

$$d^3\left(\frac{2n^3 + 3n^2 + n}{6}\right). \qquad (2)$$

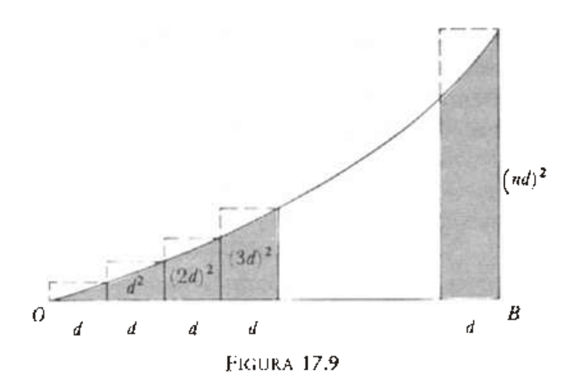


FIGURA 17.9

Pero d es la longitud fija OB dividida por n. Por tanto, (2) resulta ser

$$OB^3\left(\frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}\right). \tag{3}$$

Si se considera, como lo hicieron ellos entonces, que los dos últimos términos se pueden despreciar cuando n es infinito, se obtiene el resultado correcto. El proceso de paso al límite no había sido introducido todavía —o era percibido sólo toscamente— y por lo tanto el despreciar términos tales como los dos últimos no estaba justificado.

Vemos que el método requiere aproximar la figura curvilínea mediante otras rectilíneas, como en el método exhaustivo. Sin embargo, hay un cambio vital en el último paso: en lugar de la demostración indirecta utilizada en el método anterior, aquí el número de rectángulos se hace infinito y se toma el límite de (3) cuando n se hace infinito aunque pensar en términos de límite no era en absoluto explícito en esta época. Este nuevo enfoque, utilizado en fecha tan temprana como 1586 por Stevin en su *Statics*, fue seguido por muchos otros, incluyendo a Fermat [3].

Si la curva en cuestión no fuera la parábola, se tendría que sustituir la propiedad característica de la parábola por la de la curva en cuestión y obtener así alguna otra serie en lugar de la (1) de antes. Sumar la análoga de (1) para obtener la análoga de (2) requería ingenio. Por tanto, los resultados sobre áreas, volúmenes y centros de

---

[3] *Oeuvres*, 1, 255-259; 3, 216-219.

gravedad fueron limitados. Es claro que el potente método de calcular el límite de tales sumas invirtiendo la diferenciación todavía no había sido considerado.

Utilizando esencialmente el tipo de técnica de sumación que acabamos de ilustrar, Fermat conocía antes de 1636 que (en nuestra notación)

$$\int_0^a x^n dx = \frac{a^{n+1}}{n+1}$$

para todo n racional excepto -1 [4]. Este resultado también fue obtenido independientemente por Roberval, Torricelli y Cavalieri, aunque en algunos casos sólo en forma geométrica y para n más limitado.

Pascal fue uno de los que realizaron la sumación en forma geométrica. En 1658 consideró algunos problemas sobre la cicloide [5]. Calculó el área de cualquier segmento de la curva cortada por una recta paralela a la base, el centroide del segmento y los volúmenes de los sólidos generados por esos segmentos al girar alrededor de sus bases (YZ en la fig. 17.10) o de una recta vertical (el eje de simetría). En este trabajo, así como en trabajos previos sobre áreas encerradas bajo curvas de la familia $y = x^n$, sumó pequeños rectángulos en la forma descrita en conexión con la (1) de antes, aunque su trabajo y resultados fueron enunciados geométricamente. Bajo el seudónimo de Dettonville, proponía los problemas que había resuelto como un reto para otros matemáticos, publicando a continuación sus propias soluciones superiores (*Lettres de Dettonville*, 1659).

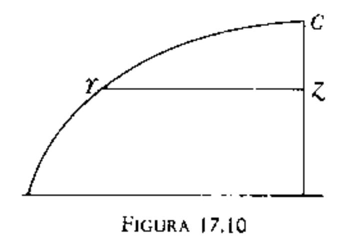


FIGURA 17.10

[4] *Oeuvres*, 1, 255-259; 3, 216-219.

[5] *Traité des sinus du quart de cercle*, 1659 = *Oeuvres*, 9, 60-76.

Antes de Newton y Leibniz, quien más hizo para introducir los métodos analíticos en el cálculo fue John Wallis (1616-1703). Aunque no comenzó a aprender matemáticas hasta que tenía aproximadamente veinte años —su educación universitaria en Cambridge estuvo dedicada a la teología— llegó a ser profesor de geometría en Oxford y el matemático británico más capaz del siglo, después de Newton. En su *Arithmetica Infinitorum* (1655) aplicó el análisis y el método de los indivisibles para efectuar muchas cuadraturas y obtener resultados amplios y útiles.

Uno de los notables resultados de Wallis, obtenido en sus esfuerzos por calcular el área del círculo analíticamente, fue una nueva expresión de π. Calculó el área acotada por los ejes, la ordenada en x y la curva para las funciones

$$y = (1 - x^2)^0, y = (1 - x^2)^1, y = (1 - x^2)^2, y = (1 - x^2)^3, \ldots$$

y obtuvo las áreas

$$x, x - \frac{1}{3}x^3, x - \frac{2}{3}x^3 + \frac{1}{5}x^5, x - \frac{3}{3}x^3 + \frac{3}{5}x^5 - \frac{1}{7}x^7, \ldots$$

respectivamente. Cuando $x = 1$, estas áreas son

$$1, \frac{2}{3}, \frac{8}{15}, \frac{48}{105}, \ldots \qquad (4)$$

Ahora bien, la circunferencia viene dada por $y = (1 - x^2)^{1/2}$. Por inducción e interpolación, Wallis calculó su área y, mediante complicados razonamientos posteriores llegó a que

$$\frac{\pi}{2} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdot 8 \ldots}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdot 9 \ldots}.$$

Gregorio de St. Vincent, en su *Opus Geometricum* (1647), proporcionó las bases para la importante conexión entre la hipérbola rectangular y la función logaritmo. Demostró, utilizando el método exhaustivo, que si para la curva $y = 1/x$ (fig. 17.11) las x_i se eligen de modo que las áreas *a, b, c, d,* ... son iguales, entonces las y_i están en progresión geométrica. Esto significa que la suma de las áreas desde x_0 hasta x_i, cuya suma forma una progresión geométrica, es

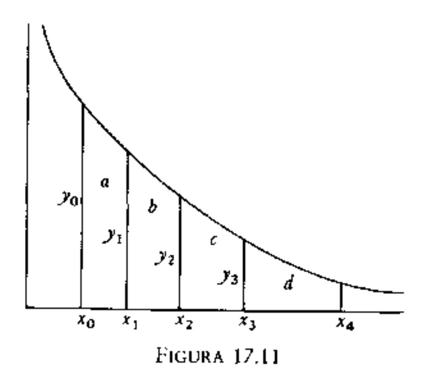


FIGURA 17.11

proporcional al logaritmo de los valores de las y_i o, en nuestra notación,

$$\int_{x_0}^{x} \frac{dx}{x} = k \log y.$$

Esto concuerda con nuestro conocido resultado del cálculo, porque $y = 1/x$. La observación de que las áreas pueden interpretarse como logaritmos se deben en realidad a un discípulo de Gregorio, el jesuita belga Alfons A. de Sarasa (1618-1667), en sus *Solutio Problematis a Mersenno Propositi* (1649). Alrededor de 1665 Newton también se dio cuenta de la conexión entre el área encerrada bajo la hipérbola y los logaritmos, e incluyó esta relación en su *Method of Fluxions*. Desarrolló $1/(1 + x)$ por el teorema del binomio e integró término a término obtenido

$$\log_e (1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$$

Nicholas Mercator, utilizando los resultados de Gregorio, obtuvo la misma serie independientemente (aunque no lo afirmó explícitamente) en su *Logarithmotechnia* de 1668. Pronto otros encontraron series que, como diríamos nosotros, convergían más rápidamente. El trabajo sobre la cuadratura de la hipérbola y su relación con la función logarítmica fue realizado por muchos, y la mayor parte se difundió por medio de cartas, por lo que es difícil detectar el orden de su descubrimiento y atribuir adecuadamente el mérito de ello.

Hasta, aproximadamente, 1650 nadie creía que la longitud de una curva pudiera ser igual exactamente a la longitud de una recta. De hecho, en el segundo libro de *La Géometrie*, Descartes expone que la relación entre las líneas curvas y rectas ni se conoce ni se podrá conocer nunca. Pero Roberval encontró la longitud de un arco de cicloide. El arquitecto Christopher Wren (1632-1723) rectificó la cicloide mostrando que (fig. 17.12) el arco $PA = 2PT$[6]. William Neile (1637-1670) también obtuvo (1659) la longitud de un arco y, utilizando una sugerencia de Wallis, rectificó la parábola semicúbica ($y^3 = ax^2$)[7]. Fermat también calculó algunas longitudes de curvas. Todos utilizaban, habitualmente, un polígono inscrito para aproximar la curva, hallaban la suma de los segmentos y entonces hacían que el número de segmentos se hiciera infinito y cada uno de ellos más pequeño. James Gregory (1638-1675), un profesor de St. Andrews y Edinburgo (cuyo trabajo fue conocido sólo ligeramente por sus contemporáneos, pero no de forma general hasta que apareció en 1939 un volumen conmemorativo, editado por H. W. Turnbull), dio un método para rectificar curvas en su *Geometriae Pars Universalis* (1668).

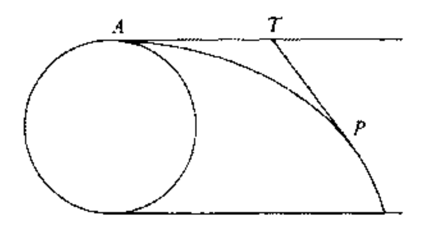


FIGURA 17.12

Christian Huygens (1629-1695) obtuvo algunos resultados adicionales sobre rectificación. En particular, encontró la longitud del arco de la cisoide. También realizó contribuciones al trabajo sobre áreas y volúmenes y fue el primero en dar resultados sobre áreas de superficies diferentes de la esfera. Así, obtuvo las áreas de porciones de superficies del paraboloide y del hiperboloide. Huygens obtuvo

---

[6] El método fue publicado por Wallis en *Tractatus Duo* (1659 = *Opera*, 1, 550-569). Wren dio sólo el resultado.

[7] El trabajo de Neile fue publicado por Wallis en la referencia de la nota 6.

todos estos resultados por métodos puramente geométricos, aunque utilizó la aritmética, como Arquímedes había hecho ocasionalmente, para obtener respuestas cuantitativas.

La rectificación de la elipse desafiaba a los matemáticos. De hecho, James Gregory afirmaba que la rectificación de la elipse y de la hipérbola no podría realizarse en términos de funciones conocidas. Durante algún tiempo, los matemáticos no se animaron a trabajar más sobre este problema, y no se obtuvieron resultados nuevos hasta el siglo siguiente.

Hemos estado tratando de las principales contribuciones de los predecesores de Newton y Leibniz a los cuatro problemas más importantes que motivaron los trabajos sobre el cálculo. Los cuatro problemas habían sido considerados como diferentes; sin embargo, se detectaron relaciones entre ellos que, incluso, llegaron a utilizarse. Por ejemplo, Fermat había usado exactamente el mismo método para obtener tangentes y para obtener el valor máximo de una función. También se vio fácilmente que el problema del cambio relativo de una función con respecto a la variable independiente y el problema de la tangente eran el mismo. De hecho, el método de Fermat y de Barrow para obtener tangentes no es más que la contrapartida geométrica de la obtención del cambio relativo. Pero la característica principal del cálculo, después de los mismísimos conceptos de derivada y de integral como límite de una suma, es el hecho de que la integral puede obtenerse invirtiendo el proceso de diferenciación o, como decimos nosotros, obteniendo la antiderivada. Se habían encontrado muchas pruebas de esta relación, pero no se había valorado adecuadamente su significado. Torricelli había observado en casos particulares que el problema del cambio relativo era esencialmente el inverso del problema del área. Estaba, de hecho, incluido en el uso que Galileo hacía del hecho de que el área encerrada bajo la curva tiempo-velocidad proporciona la distancia recorrida hasta el tiempo correspondiente. Puesto que el cambio relativo de la distancia debe de ser la velocidad, el cambio relativo del área, considerada como una «suma», debe de ser la derivada de la función de área. Pero Torricelli no cayó en la cuenta de ello en general. También Fermat conocía la relación entre área y derivada en casos particulares, pero no valoró su generalidad o importancia. James Gregory, en su *Geometriae* de 1668, demostró que los problemas de la tangente y del área son problemas inversos, pero su libro pasó desapercibido. En las *Geometrical Lectures,* Barrow exponía la relación

entre obtener la tangente a una curva y el problema del área, pero estaba dicha en forma geométrica y ni siquiera él mismo se dio cuenta de su significado.

En realidad se había acumulado una inmensa cantidad de conocimiento sobre el cálculo antes de que Newton y Leibniz entraran en escena. Una panorámica de tan sólo el único libro de Barrow muestra un método para obtener tangentes, teoremas sobre diferenciación del producto y del cociente de dos funciones, la diferenciación de potencias de x, la rectificación de curvas, el cambio de variables en una integral definida e incluso la diferenciación de las funciones implícitas. Aunque en el caso de Barrow la formulación geométrica hacía difícil el reconocimiento de las ideas generales, en la *Arithmetica Infinitorum* de Wallis podian encontrarse resultados comparables en forma aritmética.

Puede uno preguntarse entonces qué quedaba por realizar en la senda de los principales resultados nuevos. La respuesta es una mayor generalidad del método y el tomar conciencia de la generalidad de lo que ya había sido establecido en problemas particulares. Los trabajos sobre el cálculo durante los primeros dos tercios de siglo se perdieron en los detalles. Además, en sus esfuerzos por alcanzar rigor a través de la geometría, no se utilizaron ni se exploraron, en general, las implicaciones de la nueva álgebra ni de la geometría de coordenadas, y muchos se agotaron en sutiles razonamientos sin salida. Lo que estimuló, en último extremo, la percepción necesaria y el alcance de la generalidad fue el trabajo aritmético de Fermat, Gregorio de St. Vincent y Wallis, aquel a quien Hobbes criticó por sustituir la geometría por símbolos. James Gregory afirmaba en su prefacio a la *Geometriae* que la verdadera división de las matemáticas no era en geometría y aritmética, sino en lo universal y lo particular. Lo universal fue proporcionado por las dos mentes universales, Newton y Leibniz.

3. La obra de Newton

Los grandes avances en las matemáticas y en la ciencia se construyen casi siempre sobre el trabajo de muchos hombres que aportan sus contribuciones, poco a poco, a lo largo de cientos de años; de vez en cuando un hombre, lo bastante lúcido como para distinguir las ideas valiosas de sus predecesores de la confusión de sugerencias

y pronunciamientos, lo suficientemente imaginativo como para encajar las piezas en una nueva explicación, lo bastante audaz como para construir un plan maestro, da el paso culminante y definitivo. En el caso del cálculo, éste fue Isaac Newton.

Newton (1642-1727) nació en la aldea de Woolsthorpe, Inglaterra, donde su madre trabajaba la granja que le había dejado su marido, quien había muerto dos meses antes de que naciera Isaac. Se educó en escuelas locales, de bajo nivel educativo, como un joven sin ninguna inclinación especial, excepto su interés por los aparatos mecánicos. Superó los exámenes de entrada, con deficiencias en geometría euclídea, en el Trinity College de la Universidad de Cambridge en 1661, y allí estudió tranquilamente y sin obstáculos. En un determinado momento casi cambió su carrera de filosofía natural (ciencia) a derecho. Según parece, recibía muy poco estímulo de sus profesores excepto, posiblemente, de Barrow. Por ello, realizó experiencias por sí mismo y estudió la *Géometrie* de Descartes, así como los trabajos de Copérnico, Kepler, Galileo, Wallis y Barrow.

Justo después de que Newton terminara su trabajo de licenciatura, la universidad fue cerrada porque la peste se había extendido por toda el área de Londres. Se fue de Cambridge y pasó los años 1665 y 1666 en la tranquilidad de la casa familiar en Woolsthorpe. Allí inició su magnífico trabajo en mecánica, matemáticas y óptica. En esa época se dio cuenta de que la ley de gravitación del inverso del cuadrado, un concepto ya anticipado por otros, incluso por Kepler ya en 1612, era la llave de una ciencia unificadora de la mecánica; obtuvo un método general para tratar los problemas del cálculo y, mediante experimentos con la luz, realizó el trascendental descubrimiento de que la luz blanca, como la del Sol, está compuesta en realidad de todos los colores, desde el violeta hasta el rojo. «Todo esto», decía Newton más tarde, «fue en los años de la peste de 1665 y 1666, ya que en aquellos días yo estaba en lo mejor de mi edad para la invención, y me interesaban las matemáticas y la filosofía (ciencia) más que en cualquier otra época desde entonces».

Newton no dijo nada de esos descubrimientos. Volvió a Cambridge en 1667 para obtener un grado de maestro y fue elegido miembro del Trinity College. En 1669 Isaac Barrow renunció a su plaza de profesor y Newton fue contratado en el puesto de Barrow como profesor Lucasiano de matemáticas. Según parece, no era un buen profesor, porque pocos estudiantes asistían a sus clases; tampoco sus colegas se dieron cuenta de la originalidad del material que

presentaba. Sólo Barrow y, algo más tarde, el astrónomo Edmund Halley (1656-1742) valoraron su grandeza y le estimularon.

Al principio Newton no publicó sus descubrimientos. Se dice que tenía un miedo a la crítica anormal; De Morgan dice que «un miedo mórbido a la oposición de los demás gobernó toda su vida». Cuando en 1672 publicó su trabajo sobre la luz, acompañado de su filosofía de la ciencia, fue criticado severamente por la mayor parte de sus contemporáneos, incluyendo a Robert Hooke y a Huygens, quienes tenían ideas diferentes sobre la naturaleza de la luz. Newton quedó tan desconcertado que resolvió no publicar ya en el futuro. Sin embargo, en 1675 publicó otro trabajo sobre la luz que contenía su idea de que la luz era una corriente de partículas —la teoría corpuscular de la luz—. Otra vez se vio envuelto en una tormenta de críticas e incluso de pretensiones por parte de otros de haber descubierto ya esas ideas. Esta vez Newton decidió que sus resultados serían publicados después de su muerte. No obstante, publicó trabajos posteriores y varios libros famosos, los *Principia*, la *Opticks* (edición inglesa, 1704, edición italiana, 1706), y la *Arithmetica Universalis* (1707).

Desde 1665 en adelante aplicó la ley de la gravitación al movimiento planetario; en este campo, los trabajos de Hooke y Huygens le influyeron considerablemente. En 1684, su amigo Halley le instó a publicar sus resultados, pero, además de su renuncia a publicar, Newton no disponía de una demostración de que la atracción gravitatoria ejercida por una esfera sólida actúa como si la masa de la esfera estuviera concentrada en el centro. Dice, en una carta a Halley del 20 de junio de 1686, que hasta 1685 sospechaba que esto era falso. En ese año demostró que una esfera cuya densidad varía sólo con la distancia al centro atrae, de hecho, a una partícula externa como si la masa de la esfera estuviera concentrada en su centro, y se mostraba de acuerdo en escribir su trabajo.

Halley, entonces, apoyó editorialmente a Newton y pagó la publicación. En 1687 apareció la primera edición de los *Philosophiae Naturalis Principia Mathematica (Los principios matemáticos de la filosofía natural)*. Hubo dos ediciones posteriores, en 1713 y 1726, y la segunda incluía algunas mejoras. Aunque el libro le reportó a Newton una gran fama, era muy difícil de entender. Confesó a un amigo que lo había hecho difícil a propósito «para evitar ser atacado por pequeños charlatanes en matemáticas». Sin duda esperaba así evitar las críticas que habían recibido sus trabajos anteriores sobre la luz.

Newton era también un químico importante. Aunque no hay grandes descubrimientos asociados a su nombre en este campo, se debe tener en cuenta que la química estaba entonces en su infancia. Tuvo la idea correcta de intentar explicar los fenómenos químicos en términos de partículas últimas, y poseyó un conocimiento profundo de la química experimental. En este tema escribió un importante artículo, «De natura acidorum» (escrito en 1692 y publicado en 1710). En las Philosophical Transactions of the Royal Society de 1701, publicó un artículo sobre el calor, que contiene su famosa ley de enfriamiento. Aunque leía los trabajos de los alquimistas, no aceptaba sus nebulosos y místicos puntos de vista. Las propiedades químicas y físicas de los cuerpos podrían, según creía, explicarse en términos del tamaño, forma y movimiento de las partículas últimas; rechazaba las fuerzas ocultas de los alquimistas, tales como simpatía, antipatía, armonía y atracción.

Además de su obra sobre mecánica celeste, luz y química, Newton trabajó en hidrostática e hidrodinámica. A su magnífico trabajo experimental sobre la luz, añadió su experimentación sobre el rozamiento en el movimiento del péndulo en distintos medios, la caída de esferas en aire y agua, y el flujo de agua de surtidores. Como la mayor parte de los científicos de la época, Newton se construyó su propio equipo. Se construyó dos telescopios reflectantes, produciendo incluso la aleación para el armazón, fabricando la montura y puliendo las lentes.

Después de trabajar como profesor durante treinta y nueve años, Newton se deprimió y sufrió un colapso nervioso. Decidió abandonar la investigación y en 1695 aceptó un contrato como director de la Casa de la Moneda de Londres. Durante sus veintisiete años en ese puesto no hizo investigación, salvo para trabajar ocasionalmente en algún problema. Llegó a ser presidente de la Royal Society en 1703, cargo que mantuvo hasta su muerte; fue hecho caballero en 1705.

Es evidente que Newton se interesó mucho más en la ciencia en general que en las matemáticas, y que participó activamente en los problemas de su época. Consideró que el principal valor de su trabajo científico era el de que constituyera un apoyo de la religión revelada y era, de hecho, un teólogo ilustrado, aunque nunca se ordenó. Pensaba que la investigación científica era dura y monótona, pero se mantuvo en ella porque proporcionaba pruebas de la gran obra de Dios. Como su antecesor Barrow, Newton se orientó hacia

los estudios religiosos relativamente tarde. En la obra *The Chronology of Ancient Kingdoms Amended*, intentó atribuir fecha precisa a sucesos descritos en la Biblia y en otros documentos religiosos, relacionándolos con sucesos astronómicos. Su principal trabajo religioso fue las *Observations Upon the Prophecies of Daniel and the Apocalypse of St. John*. La exégesis biblica fue una fase del enfoque racional de la religión, que era común en la Edad de la Razón; Leibniz también intervino en ello.

Al menos por lo que se refiere al cálculo, Newton generalizó las ideas ya adelantadas por muchos otros, estableció métodos ya maduros y mostró las interrelaciones entre varios de los importantes problemas descritos anteriormente. Aunque aprendió mucho como alumno de Barrow, en álgebra y cálculo estuvo más influido por los trabajos de Wallis. Decía que se vio conducido a sus descubrimientos en análisis por la *Arithmetica Infinitorum;* ciertamente, realizó progresos en su trabajo sobre el cálculo razonando analíticamente. Sin embargo, incluso Newton pensaba que la geometría era necesaria para una demostración rigurosa.

En 1669 Newton hizo circular entre sus amigos una monografía titulada *De Analysi per Aequationes Numero Terminorum Infinitas* (Sobre el análisis por medio de ecuaciones con un infinito número de términos); no fue publicado hasta 1711. Supone que tiene una curva, y que el área z (fig. 17.13) bajo esa curva viene dada por

$$z = ax^m, \tag{5}$$

donde m es entero o fraccionario. A un incremento infinitesimal de x lo llama momento de x, y lo representa mediante o, una notación utilizada por James Gregory y que equivale a la E de Fermat. Al área acotada por la curva, el eje OX, el eje OY y la ordenada en $x + o$, la llama $z + oy$, donde oy es el momento del área. Entonces

$$z + oy = a(x + o)^m. \tag{6}$$

Aplica el teorema del binomio al segundo miembro, obteniendo una serie infinita cuando m es fraccionario, resta (5) de (6), divide por o, desprecia los términos que contienen todavía o y obtiene

$$y = max^{m-1}.$$

Así, en el lenguaje actual, el cambio relativo del área en cualquier x

es el valor de y de la curva en ese valor de x. Recíprocamente, si la curva es $y = max^{m-1}$, el área encerrada por ella es $z = ax^m$.

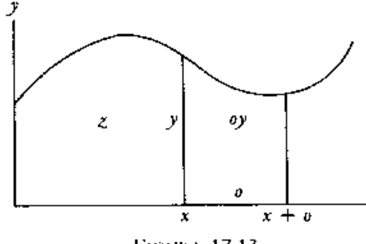


FIGURA 17.13

En este proceso Newton no sólo dio un método general para obtener el cambio relativo de una variable con respecto a otra (z con respecto a x en el ejemplo anterior), sino que mostró que el área puede obtenerse invirtiendo el proceso de obtener un cambio relativo. Como las áreas se habían obtenido y expresado también como sumación de áreas infinitesimales, Newton también mostró que tales sumas pueden obtenerse mediante la inversión del proceso de obtener un cambio relativo. Este hecho, el de que las sumaciones (más propiamente, los límites de sumas) puedan obtenerse invirtiendo la diferenciación, es lo que llamamos ahora el teorema fundamental del cálculo. Aunque era conocido en casos especiales y confusamente previsto por los predecesores de Newton, él vio que era general. Aplicó el método para obtener el área encerrada bajo muchas curvas y para resolver otros problemas que pueden resolverse como sumaciones.

Después de demostrar que la derivada del área es el valor de la y y afirmar que el recíproco es cierto, Newton estableció la regla de que, si el valor de y es una suma de términos, entonces el área es la suma de las áreas que resultan de cada uno de los términos. En expresión moderna, la integral indefinida de una suma de funciones es la suma de las integrales de las funciones por separado.

En su siguiente contribución en la monografía citada, desarrolló su utilización de las series infinitas. Para integrar $y = a^2/(b + x)$, dividió a^2 por $b + x$ y obtuvo

$$y = \frac{a^2}{b} - \frac{a^2x}{b^2} + \frac{a^2x^2}{b^3} - \frac{a^2x^3}{b^4} + \ldots$$

Habiendo obtenido esta serie infinita, calcula la integral integrando término a término, de forma que el área es

$$\frac{a^2x}{b} - \frac{a^2x^2}{2b^2} + \frac{a^2x^3}{3b^3} - \frac{a^2x^4}{4b^4} + \dots$$

Dice de esta serie infinita que unos pocos de los primeros términos son suficientemente exactos para cualquier uso, con tal que b sea igual a x repetido algunas veces.

De la misma manera, para integrar $y = 1/(1 + x^2)$ utiliza el desarrollo del binomino para escribir

$$y = 1 - x^2 + x^4 - x^6 + x^8 - \dots$$

e integra término a término. Hace notar que si, en lugar de ello, se toma $y = 1/(1 + x^2)$, mediante el desarrollo del binomio se obtendría

$$y = x^{-2} - x^{-4} + x^{-6} - x^{-8} + \dots$$

que se puede integrar término a término. Señala entonces que cuando x es suficientemente pequeño debe utilizarse el primer desarrollo, pero que cuando x es grande debe usarse el segundo. Por tanto, de alguna manera era consciente de que lo que llamamos convergencia era importante, pero no tenía una noción precisa de ello.

Newton se dio cuenta de que había extendido la integración término a término a las series, pero dice en *De Analysi:*

> Y cualquier cosa que realice el Análisis común mediante Ecuaciones de un Número de Términos finito (con tal que se pueda hacer) se puede hacer siempre mediante Ecuaciones infinitas, de modo que no he tenido ningún reparo en llamar a esto Análisis igualmente. Porque los razonamientos en este campo no son menos ciertos que en el otro; tampoco las ecuaciones menos exactas; aunque nosotros Mortales, cuyos poderes de razonamiento están confinados dentro de límites estrechos, no podemos ni expresar ni tampoco concebir todos los Términos de estas Ecuaciones, como para conocer exactamente de ellos las cantidades que queremos.

Hasta entonces, en ese enfoque del cálculo, Newton había utilizado lo que puede ser descrito como el método de los infinitesimales. Los momentos son cantidades infinitamente pequeñas, indivisibles o infinitesimales. La lógica de lo que hizo Newton no está clara,

por supuesto. Dice en este trabajo que su método está «explicado brevemente más que demostrado con precisión».

Newton proporcionó una segunda, más extensa y más definitiva, exposición de sus ideas en el libro *Methodus Fluxionum et Serierum Infinitarum*, escrito en 1672 pero no publicado hasta 1736. En este trabajo dice que considera sus variables como generadas por el movimiento continuo de puntos, rectas y planos, más que como agregados estáticos de elementos infinitesimales, como en el artículo anterior. Ahora, a una cantidad variable la llama fluyente, y a su cambio relativo, fluxión. Su notación es $\dot{x}$ y $\dot{y}$ y para fluxiones de x y de y. La fluxión de $\dot{x}$ es $\ddot{x}$, etc. La fluyente de la que x es la fluxión es $\acute{x}$ y la fluyente de ésta es $\acute{x}'$.

En este segundo trabajo Newton establece algo más claramente el problema fundamental del cálculo: dada una relación entre dos fluyentes, obtener la relación entre sus fluxiones y recíprocamente. Las dos variables de las que se da la relación pueden representar cualquier cantidad. Sin embargo, Newton piensa en ellas como cambiantes con el tiempo porque es una forma de pensar útil, aunque, como señala, no es necesaria. Por tanto, si o es un «intervalo de tiempo infinitamente pequeño», entonces $\dot{x}o$ y $\dot{y}o$ son los incrementos indefinidamente pequeños de x y de y, o los momentos de x e y. Para obtener la relación entre $\dot{x}$ e $\dot{y}$, supongamos, por ejemplo, que la fluyente es $y = x^n$. Newton forma, primero,

$$y + \dot{y}o = (x + \dot{x}o)^n,$$

y entonces procede como en el artículo anterior. Desarrolla el segundo miembro mediante el teorema del binomio, resta $y = x^n$, divide por o, desprecia todos los términos que todavía contienen o y obtiene

$$\dot{y} = nx^{n-1}\dot{x}.$$

En notación moderna, este resultado puede escribirse

$$\frac{dy}{dt} = nx^{n-1}\,\frac{dx}{dt},$$

y como $dy/dx = (dy/dt)/(dx/dt)$, Newton, al obtener el cociente de dy/dt y dx/dt o de $\dot{y}$ y $\dot{x}$, obtuvo dy/dx.

El método de las fluxiones no es esencialmente diferente del utilizado en *De Analysi,* y tampoco el rigor es mayor; Newton desprecia términos como $\dot{x}\dot{x}o$ y $\dot{x}\dot{x}o\dot{x}o$ (él escribe $\dot{x}^3oo$) sobre la base de que son infinitamente pequeños comparados con el que se conserva. Sin embargo, su punto de vista en el *Method of Fluxions* es algo diferente. Los momentos $\dot{x}o$ e $\dot{y}o$ cambian con el tiempo o, mientras que en el primer artículo los momentos son en realidad trozos fijos de x y z. Esta visión, más nueva, sigue el pensamiento más dinámico de Galileo; la más antigua utilizaba el indivisible estático de Cavalieri. El cambio servía, según decía Newton, sólo para eliminar la aspereza de la doctrina de los indivisibles; sin embargo, los momentos $\dot{x}o$ y $\dot{y}o$ son todavía algún tipo de cantidades infinitamente pequeñas. Más todavía, $\dot{x}$ e $\dot{y}$, que son las fluxiones o derivadas con respecto al tiempo de x e y, en realidad no están nunca definidas; este problema central se soslaya.

Dada una relación entre $\dot{x}$ e $\dot{y}$, obtener la relación entre x e y es más difícil que la mera integración de una función de x. Newton trata varios tipos: (1) cuando $\dot{x}$, $\dot{y}$, y x o y están presentes; (2) cuando $\dot{x}$, $\dot{y}$, x e y están presentes; (3) cuando $\dot{x}$, $\dot{y}$, $\dot{z}$, y las fluyentes están presentes. El primer tipo es el más fácil y, en notación moderna, corresponde a resolver $dy/dx = f(x)$. Del segundo tipo, Newton trata $\dot{y}/\dot{x} = 1 - 3x + y + x^2 + xy$ y la resuelve mediante un proceso de aproximaciones sucesivas. Comienza con $\dot{y}/\dot{x} = 1 - 3x + x^2$ como primera aproximación, obtiene y como función de x, introduce este valor de y en el segundo miembro de la ecuación original, y continúa el proceso. Newton describe lo que hace, pero no lo justifica. Del tercer tipo trata $2\dot{x} - \dot{z} + \dot{y}x = 0$. Supone una relación entre x e y, por ejemplo, $x = y^2$, de modo que $\dot{x} = 2\dot{y}y$. Entonces la ecuación se convierte en $4\dot{y}y - \dot{z} - \dot{y}y^2 = 0$, de la que obtiene $2y^2 + (y^3/3) = z$. Por tanto, si el tercer tipo se considera una ecuación en derivadas parciales, Newton obtiene sólo una integral particular.

Newton se dio cuenta de que en este trabajo había presentado un método general. En una carta a John Collins, fechada el 10 de diciembre de 1672, en la que proporciona los elementos de su método y un ejemplo, dice,

> Este es un [caso] particular, o más bien un corolario, de un método general, que puede aplicarse, sin ningún cálculo complicado, no sólo al dibujo de las tangentes de cualquier línea curva, tanto geométrica como mecánica... sino también para resolver otros tipos más abstrusos de problemas sobre curva-

turas, áreas, longitudes, centros de gravedad de curvas, etc.; tampoco está... limitado a las ecuaciones que no contengan cantidades irracionales. He entretejido este método con el de las ecuaciones, reduciéndolos a las series infinitas.

Newton resaltaba el uso de las series infinitas porque mediante ellas podía tratar funciones tales como $(1 + x)^{3/2}$, mientras que sus predecesores habían estado limitados, en su conjunto, a las funciones algebraicas racionales.

En su *Tractatus de Quadratura Curvarum* (Tratado sobre la cuadratura de las curvas), un tercer artículo sobre el cálculo, escrito en 1676 pero publicado en 1704, Newton dice que había abandonado el infinitesimal o cantidad infinitamente pequeña. Critica ahora el despreciar los términos que incluyen o porque, según dice

en matemáticas no se debe despreciar ni los errores más diminutos... Considero las cantidades matemáticas, en este punto, no como consistentes en pequeñas partes, sino como descritas por un movimiento continuo. Las líneas están descritas, y por tanto generadas, no por la yuxtaposición de partes, sino por el movimiento continuo de puntos; las superficies por el movimiento de líneas; los ángulos por la rotación de los lados; las porciones de tiempo por un flujo continuo...

Las fluxiones son, hasta la aproximación que queramos, como los incrementos de las fluyentes generados en tiempos iguales y tan pequeños como sea posible y, para hablar con precisión, están en la razón primera de los incrementos emergentes; aunque pueden expresarse mediante líneas cualesquiera que sean proporcionales a ellos.

El nuevo concepto de Newton, el método de la razón primera y última, significa lo siguiente. Considera la función $y = x^n$. Para obtener la fluxión de y o de x^n, se deja a x «fluir» hasta $x + o$. Entonces x^n se convierte en

$$(x + o)^n = x^n + nox^{n-1} + \frac{n^2 - n}{2} o^2x^{n-2} + \ldots$$

Los incrementos de x e y, es decir, o y $nox^{n-1} + \frac{n^2 - n}{2} o^2x^{n-2} + \ldots$ son, uno con respecto al otro (dividiendo ambos por o), como

$$1 \text{ a } nx^{n-1} + \frac{n^2 - n}{2} ox^{n-2} + \dots$$

«Hagamos ahora tender a cero los incrementos, y la última proporción se convertirá en

$$1 \text{ a } nx^{n-1}.$$

Entonces, la fluxión de x es a la fluxión de x^n como 1 a nx^{n-1} o, como diríamos en la actualidad, el cambio relativo de y con respecto a x es nx^{n-1}. Esta es la razón o cociente primera de los incrementos emergentes. Por supuesto que la lógica de esta versión no es mejor que la de las dos precedentes; sin embargo, Newton dice que su método está en armonía con la geometría de los antiguos y que no es necesario introducir cantidades infinitamente pequeñas.

Newton también dio una interpretación geométrica. Dados los datos de la figura 17.14, supongamos que bc se mueve hasta BC de modo que c coincida con C. Entonces el triángulo curvilíneo CEc es «en la última forma» semejante al triángulo CET, y sus lados «evanescentes» serán proporcionales a CE, ET y CT. Por lo tanto, las fluxiones de las cantidades AB, BC y AC son, en la última razón o cociente de sus incrementos evanescentes, proporcionales a los lados del triángulo CET o del VBC.

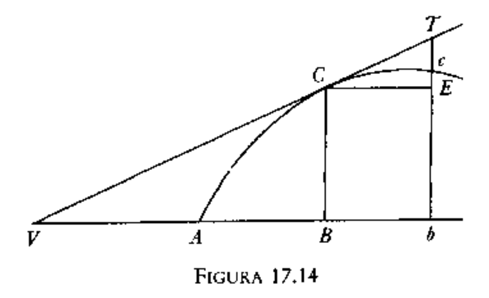


FIGURA 17.14

En el *Method of Fluxions* Newton incluyó aplicaciones las fluxiones a la diferenciación de las funciones implícitas y a la obtención de tangentes a las curvas, máximos y mínimos de las funciones, curvatura de las curvas y puntos de inflexión de las mismas. También obtuvo áreas y longitudes de algunas curvas. En conexión con la curvatura, dio la fórmula correcta del radio de curvatura, es decir,

$$r = \frac{(1 + \dot{y}^2)^{3/2}}{\ddot{y}}$$

donde $\dot{x}$ se toma como 1. Dio también la misma cantidad en coordenadas polares. Finalmente, incluyó una breve tabla de integrales.

Newton no publicó sus artículos básicos sobre el cálculo hasta mucho después de haberlos escrito. El primer informe impreso de su teoría de las fluxiones apareció en el *Algebra* de Wallis (segunda edición en latín, 1693), de la que Newton escribió desde la página 390 hasta la 396. Si la hubiera publicado entonces de una vez, podría haber evitado la controversia con Leibniz sobre la prioridad del descubrimiento.

La primera publicación de Newton que incluye su desarrollo del cálculo es la magnífica obra *Mathematical Principles of Natural Philosophy* [8]. Por lo que se refiere a la noción básica del cálculo, la fluxión o, como diríamos nosotros, la derivada, Newton hace varias afirmaciones. Rechaza los infinitesimales o las cantidades indivisibles últimas en favor de las «cantidades divisibles evanescentes», cantidades que se puede hacer disminuir tanto como se quiera. En las ediciones primera y tercera de los *Principia* Newton dice: «Cocientes (o razones) últimos en los que las cantidades se anulan no son, estrictamente hablando, razones de cantidades últimas sino límites a los que se acercan las razones de esas cantidades, al decrecer sin límite, las cuales, aunque pueden hacerse más próximos (a sus límites) que cualquier diferencia dada, no pueden ni sobrepasarlos ni alcanzarlos antes que las cantidades hayan decrecido indefinidamente» [9]. Esta es la afirmación más clara de todas las que hizo con respecto a su cociente último. A propósito de la cita anterior, también dice: «Por velocidad última se entiende aquella con la que se mueve el cuerpo, ni antes de que llegue a su posición final, cuando cesa el movimiento, ni después, sino en el mismo instante en el que llega... Y, de la misma manera, por razón última de cantidades evanescentes debe entenderse la razón de cantidades, no antes de que se anulen, no después, sino aquella con la que se anulan.»

---

[8] La tercera edición fue traducida al inglés por Andrew Motte en 1729. Esta edición, revisada y editada por Florian Cajori, fue publicada por la University of California Press en 1946.

[9] Tercera edición, p. 39.

En los *Principia* Newton utilizó métodos de demostración geométricos. Sin embargo, en lo que se llaman los Portsmouth Papers (artículos de Portsmouth), que contienen trabajos no publicados, utilizó métodos analíticos para obtener algunos de los teoremas. Estos artículos muestran que él también obtenía resultados analíticamente, más allá de los que podía traducir en términos geométricos. Se cree que una razón por la cual recurría a la geometría es porque las demostraciones resultaban más comprensibles para sus contemporáneos. Otra es que admiraba inmensamente el trabajo geométrico de Huygens y esperaba igualarlo. En estas demostraciones geométricas Newton utiliza los procesos de paso al límite básicos del cálculo. Así, el área encerrada bajo una curva se considera esencialmente como la suma de los rectángulos que la aproximan, tal como se hace en el cálculo actualmente. Sin embargo, en vez de calcular tales áreas, utiliza este concepto para comparar áreas encerradas bajo diferentes curvas.

Demuestra que, cuando AR y BR (fig. 17.15) son las perpendiculares a las tangentes en A y en B al arco ACB, la razón última, cuando B se aproxima y llega a coincidir con A, de dos cualesquiera de las cantidades, la cuerda AB, el arco ACB y AD, es 1. Por ello, dice en el corolario 3 al lema 2 del libro I: «Y, en consecuencia, en toda nuestra argumentación sobre razones últimas, podemos utilizar libremente una de esas líneas en lugar de otra cualquiera de ellas.» Demuestra entonces que cuando B se acerca y llega a coincidir con A, la razón de dos triángulos cualesquiera (de sus áreas), RAB, $RACB$ y RAD es 1. «Y, por tanto, en toda la argumentación sobre razones últimas podemos utilizar uno de esos triángulos en lugar de otro cualquiera de ellos.» Además (fig. 17.16), sean BD y CE perpendiculares a AE (que no es necesariamente tangente al arco ABC en A). Cuando B y C se aproximan y coinciden con A, la razón última de las áreas ACE y ABD es igual a la razón última de AE^2 a AD^2.

Los *Principia* contienen una gran riqueza de resultados, de los que señalaremos algunos. Aunque el libro está dedicado a la mecánica celeste, tiene una enorme importancia para la historia de las matemáticas, no sólo porque el propio trabajo de Newton sobre el cálculo estuvo motivado en gran parte por su constante interés por los problemas allí tratados, sino porque los *Principia* presentaban nuevos temas y enfoques de problemas que fueron explorados durante los cien años siguientes, en el curso de los cuales fue creada una parte enorme del análisis.

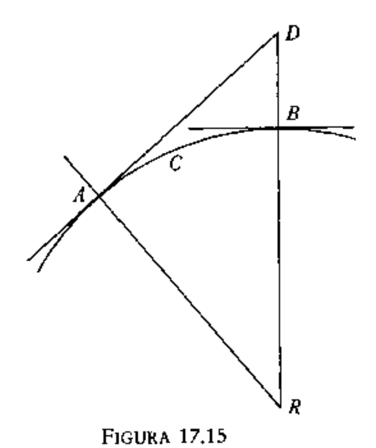


FIGURA 17.15

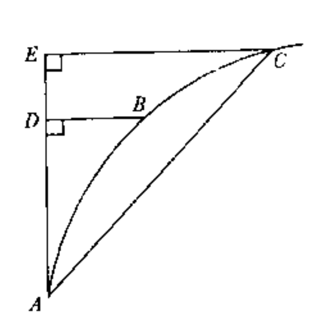


FIGURA 17.16

Los *Principia* están divididos en tres libros [10]. En una sección preliminar, Newton define conceptos de mecánica tales como los de inercia, momento y fuerza, y a continuación establece los tres famosos axiomas o leyes del movimiento. En sus palabras, son:

Ley I. Todo cuerpo continúa en su estado de reposo, o de movimiento uniforme en línea recta, a menos que se vea impelido a cambiar ese estado por fuerzas que actúen sobre él.

Ley II. El cambio [en la cantidad] de movimiento es proporcional a la potencia motriz actuante; y se realiza en la dirección de la línea recta en la que se imprime esa fuerza.

Por cantidad de movimiento Newton indica, como había explicado anteriormente, el producto de la masa por la velocidad. Por tanto, el cambio en el movimiento, si la masa es constante, es el cambio en velocidad, es decir, la aceleración. Esta segunda ley se escribe ahora a menudo como $F = ma$, cuando la fuerza F está en newtons, la masa m en kilogramos y la aceleración a en metros por segundo en cada segundo. La segunda ley de Newton es en realidad un enunciado vectorial; es decir, si la fuerza tiene componentes en, digamos, tres direcciones mutuamente perpendiculares, entonces cada componente provoca una aceleración en su propia dirección. Newton utilizó el carácter vectorial de la fuerza en problemas particulares, pero todo el significado de la naturaleza vectorial de la ley fue

[10] Todas las referencias están hechas con respecto a la edición mencionada en la nota 8.

completamente reconocido por primera vez por Euler. Esta ley incorpora el cambio clave con respecto a la mecánica de Aristóteles, la cual afirmaba que la fuerza es la causa de la velocidad. Aristóteles también había afirmado que se necesita una fuerza para mantener la velocidad. La Ley I lo niega.

Ley III. A toda acción siempre se opone una reacción igual...

No entraremos en una digresión sobre la historia de la mecánica, excepto para señalar que las dos primeras leyes son más explícitas y son enunciados, algo generalizados, de los principios del movimiento descubiertos previamente y avanzados ya por Galileo y Descartes. La distinción entre masa, esto es, la resistencia que ofrece un cuerpo a un cambio en su movimiento, y peso, la acción que la gravedad ejerce sobre la masa de cualquier objeto, también se debe a ellos; y el carácter vectorial de la fuerza generaliza el principio de Galileo de que los movimientos horizontales y verticales de un proyectil, por ejemplo, pueden tratarse independientemente.

El libro I de los *Principia* comienza con algunos teoremas del cálculo, incluyendo los que relacionan las razones últimas citadas anteriormente. Trata a continuación el movimiento bajo fuerzas centrales, es decir, fuerzas que siempre atraen al objeto móvil hacia un punto (fijo, que resulta ser el Sol, en la práctica), y demuestra en la Proposición 1 que áreas iguales son barridas en tiempos iguales (que abarca la ley de las áreas de Kepler). Newton considera a continuación el movimiento de un cuerpo a lo largo de una sección cónica y muestra (Props. 11, 12 y 13) que la fuerza debe variar como el inverso del cuadrado de la distancia a un punto fijo. También prueba el recíproco, que contiene la primera ley de Kepler. Después de algún tratamiento de la fuerza centrípeta, deduce la tercera ley de Kepler (Prop. 15). Siguen dos secciones dedicadas a las propiedades de las secciones cónicas. El problema principal es la construcción de cónicas que satisfagan cinco condiciones dadas; en la práctica éstas son, habitualmente, datos de observación. Entonces, dado el tiempo que un objeto ha estado en movimiento a lo largo de una sección cónica, determina su velocidad y posición. Estudia el movimiento de las líneas absidales, es decir, las líneas que unen el centro de atracción (en un foco) a la distancia máxima o mínima de un cuerpo que se mueve a lo largo de una cónica, que gira, a su vez, a cierta velocidad alrededor del foco. La Sección 10 está dedicada al movimiento de cuerpos sobre superficies, con referencia especial al movimiento del péndulo. Aquí Newton reconoce debidamente el trabajo de Huy-

gens. En conexión con el efecto acelerador de la gravedad sobre los movimientos, investiga las propiedades geométricas de las cicloides, epicicloides e hipocicloides, y proporciona la longitud de la epicicloide (Prop. 49).

En la Sección 11 Newton deduce, a partir de las leyes del movimiento y la ley de la gravitación, el movimiento de dos cuerpos que se atraen mutuamente con la fuerza debida a la gravitación. Su movimiento se reduce al movimiento de uno de ellos alrededor del segundo cuerpo que se toma fijo. El cuerpo que se mueve recorre una elipse.

Considera a continuación la atracción ejercida por esferas y esferoides, de densidad uniforme y variable, sobre una partícula. Da una demostración geométrica (Sec. 12, Prop. 70) de que una fina capa esférica homogénea no ejerce ninguna fuerza sobre una partícula en su interior. Como su resultado se verifica para una capa delgada, también lo hace para una suma de tales capas, esto es, para una capa de espesor finito. (Demuestra más adelante [Prop. 91, Cor. 3] que se verifica el mismo resultado para una capa elipsoidal homogénea, es decir, una capa contenida entre dos superficies elipsoidales semejantes, situadas de forma similar.) La Proposición 71 muestra que la atracción de una fina capa esférica y homogénea sobre una partícula *externa* es equivalente a la atracción que se ejercería si la masa de la capa estuviera concentrada en el centro, de modo que la capa atrae a la partícula hacia el centro y con una fuerza que varía como el inverso del cuadrado de la distancia al centro. La Proposición 73 muestra que una esfera sólida homogénea atrae a una partícula interior con una fuerza proporcional a la distancia de la partícula al centro. Por lo que se refiere a la atracción que ejerce una esfera sólida y homogénea sobre un punto externo, la Proposición 74 muestra que es la misma que si la masa de la esfera estuviera concentrada en el centro. Por lo tanto, si dos esferas se atraen mutuamente, la primera atrae a todas las partículas de la segunda como si la masa de la primera estuviera concentrada en su centro. Entonces la primera esfera se convierte en una partícula atraída por la masa distribuida de la segunda, por lo que la segunda esfera también puede tratarse como una partícula cuya masa está concentrada en su centro. Así, ambas esferas pueden tratarse como partículas cuyas masas están concentradas en sus centros respectivos. Todos estos resultados, originales de Newton, se extienden a esferas cuyas densidades son simétricas esféricamente, así como a otras leyes de atracción además de la del inverso del cuadrado de la distancia.

A continuación Newton considera el movimiento de tres cuerpos, cada uno de los cuales atrae a los otros dos, y obtiene algunos resultados aproximados. El problema del movimiento de tres cuerpos ha sido muy importante desde la época de Newton, y hasta el momento no ha sido resuelto exactamente.

El segundo libro de los *Principia* está dedicado al movimiento de cuerpos en medios resistentes tales como el aire o los líquidos. Es el comienzo de la hidrodinámica. Newton supone, en algunos problemas, que la resistencia del medio es proporcional a la velocidad y, en otros, al cuadrado de la velocidad del cuerpo que se mueve. Considera también qué forma debe tener un cuerpo para encontrar la mínima resistencia (Cap. 24, sec. 1). Asimismo considera el movimiento de péndulos y proyectiles en el aire y en fluidos. Dedica una sección a la teoría de ondas en el aire (por ejemplo, ondas de sonido) y obtiene una fórmula para la velocidad del sonido en el aire. También trata el movimiento de ondas en el agua. Newton continúa con una descripción de los experimentos que realizó para determinar la resistencia que los fluidos ofrecen a los cuerpos que se mueven en su seno. Una conclusión importante es la de que los planetas se mueven en un vacío. En este libro, Newton abrió un campo enteramente nuevo; sin embargo, el trabajo definitivo sobre el movimiento de los fluidos estaba todavía por hacer.

El libro III, titulado *On the System of the World* (Sobre el sistema del Mundo), contiene la aplicación de la teoría general desarrollada en el libro I al sistema solar. Muestra cómo puede calcularse la masa del Sol en términos de la de la Tierra, y que la masa de cualquier planeta que tenga un satélite puede obtenerse de la misma manera. Calcula la densidad media de la Tierra y obtiene que está entre cinco y seis veces la del agua (la cifra actual está alrededor de 5,5).

Muestra que la Tierra no es una esfera verdadera sino un esferoide oblato y calcula el aplanamiento; su resultado es que la elipticidad del esferoide oblato es de 1/230 (la cifra actual es de 1/297). De la observación de esta forma en cualquier planeta, se puede calcular la longitud de su día. Utilizando el grado de aplanamiento y la noción de fuerza centrípeta, Newton calcula la variación de la atracción gravitatoria de la Tierra sobre la superficie y, por tanto, la variación en el peso de un objeto. Demuestra que la fuerza atractiva de un esferoide no es la misma que si la masa del esferoide estuviera concentrada en su centro.

Se ocupa a continuación de la precesión de los equinoccios. La explicación se basa en el hecho de que la Tierra no es esférica, sino que se abomba a lo largo del Ecuador. En consecuencia, la atracción gravitatoria de la Luna sobre la Tierra no actúa en realidad sobre el centro de la Tierra, sino que fuerza un cambio periódico en la dirección del eje de rotación de la Tierra. Newton calculó el período de este cambio, y obtuvo que era de 26.000 años, valor obtenido por Hiparco a partir de las observaciones a su disposición.

Newton explicó las principales características de las mareas (libro I, Prop. 66; libro III, Prop. 36, 37). La Luna es la causa principal; el Sol, la segunda. Utilizando la masa del Sol calculó la altura de las mareas solares. De las alturas observadas de las mareas más altas y más bajas (el Sol y la Luna en plena conjunción o en plena oposición) determinó la marea lunar e hizo una estimación de la masa de la Luna. Newton también fue capaz de efectuar un tratamiento aproximado del efecto del Sol sobre el movimiento de la Luna alrededor de la Tierra. Determinó el movimiento de la Luna en latitud y en longitud; el movimiento de la línea de los ábsides (la línea que va del centro de la Tierra al punto de máxima distancia de la Luna); el movimiento de los nodos (los puntos en los que la trayectoria de la Luna corta al plano de la órbita de la Tierra; estos puntos efectúan un movimiento de regresión, es decir, se mueven lentamente en dirección opuesta al movimiento de la Luna misma); la evección (un cambio periódico de la excentricidad de la órbita de la Luna); la ecuación anual (el efecto sobre el movimiento de la Luna del cambio diario de distancia entre la Tierra y el Sol); y el cambio periódico en la inclinación del plano de la órbita de la Luna con respecto al plano de la órbita de la Tierra. Había siete irregularidades conocidas en el movimiento de la Luna, y Newton descubrió dos más, las desigualdades del apogeo (la línea de los ábsides) y de los nodos. Sus aproximación proporcionó sólo la mitad del movimiento de la línea de los ábsides. Clairaut, en 1752, mejoró el cálculo y obtuvo los 3° de rotación de la línea de los ábsides; sin embargo, mucho más tarde, John Couch Adams halló el cálculo correcto en los papeles de Newton. Finalmente, Newton mostró que los cometas deben moverse bajo la atracción gravitatoria del Sol porque sus trayectorias, determinadas sobre la base de observaciones, son secciones cónicas. Newton dedicó una gran cantidad de tiempo al problema del movimiento de la Luna porque, como observamos en el capítulo anterior, este conocimiento era necesario para mejorar el

método para determinar la longitud. Trabajó tanto en ello que se quejaba de que le producía dolores de cabeza.

4. La obra de Leibniz

Aunque sus contribuciones fueron bastante diferentes, el hombre que se alinea con Newton en la construcción del cálculo es Gottfried Wilhelm Leibniz (1646-1716). Estudió leyes y, después de defender una tesis sobre lógica, recibió el grado de licenciado en Filosofía. En 1666 escribió la tesis *De Arte Combinatoria* (Sobre el arte de las combinaciones) [11], un trabajo sobre un método universal de razonamiento; esto completaba su trabajo para un doctorado en filosofía en la universidad de Altdorf y le cualificaba para una plaza de profesor. Durante los años 1670 y 1671 Leibniz escribió sus primeros artículos sobre mecánica y hacia 1671 había producido su máquina de calcular. Obtuvo un trabajo como embajador del Elector de Mainz y en marzo de 1672 marchó a París en una misión política. Esta visita le puso en contacto con matemáticas y científicos, en particular con Huygens, y despertó su interés en las matemáticas. Aunque había leído algo sobre el tema y había escrito el artículo de 1666, él dice que no conocía casi las matemáticas hasta 1672. En 1673 fue a Londres y conoció a otros científicos y matemáticos, entre los que se encontraba Henry Oldenburg, en aquel tiempo secretario de la Royal Society de Londres. Aunque trabajando como diplomático profundizó en las matemáticas y leyó a Descartes y a Pascal. En 1676 Leibniz fue contratado como bibliotecario y consejero del Elector de Hannover. Veintidós años más tarde, el Elector de Brandenburgo invitó a Leibniz a trabajar para él en Berlín. Aunque envuelto en todo tipo de maniobras políticas, entre ellas la de la sucesión de Georg Ludwig de Hannover al trono de Inglaterra, Leibniz trabajó en muchos temas, y sus actividades colaterales cubrieron un campo enorme. Murió despreciado en 1716.

Además de diplomático, Leibniz fue filosófo, abogado, historiador, filólogo y geólogo pionero. Realizó un trabajo importante en lógica, mecánica, óptica, matemáticas, hidrostática, neumática, ciencia náutica y máquinas de calcular. Aunque su profesión era la jurisprudencia, sus trabajos en matemáticas y filosofía están entre los

---

[11] Publicado en 1690 = *Die philosophische Schriften*, 4, 27-102.

mejores que ha producido el mundo. Mantuvo contacto por carta con gente en sitios tan alejados como China y Ceilán. Intentó incansablemente reconciliar las iglesias católica y protestante. Fue quien propuso, en 1669, que se fundara una Academia de Ciencias Alemana; al fin, la Academia de Berlín fue organizada en 1700. Su recomendación original había sido la de una sociedad para la realización de inventos en mecánica y descubrimientos en química y fisiología que pudieran ser útiles a la humanidad; Leibniz quería que el conocimiento se aplicara. Llamaba a las universidades «monacales» y las acusaba de que poseían los conocimientos, pero no la facultad de discernir, y que estaban absortas en cuestiones sin importancia. En lugar de ello, urgía la búsqueda del conocimiento real —matemáticas, física, geografía, química, anatomía, botánica, zoología e historia—. Para Leibniz, las habilidades del artesano y del hombre práctico eran más valiosas que las cultas sutilezas de los eruditos profesionales. Favoreció la lengua alemana sobre el latín porque éste era el aliado del pensamiento antiguo e inútil. Los hombres enmascaran su ignorancia, decía, utilizando la lengua latina para impresionar a la gente. El alemán, por otra parte, era entendido por la gente de la calle, y podía desarrollarse para contribuir a la claridad del pensamiento y a la agudeza del razonamiento.

Leibniz publicó artículos sobre el cálculo desde 1684 en adelante, y diremos más sobre ellos después. Sin embargo, muchos de sus resultados, así como el desarrollo de sus ideas, están contenidos en cientos de páginas de notas hechas desde 1673 en adelante, pero que él nunca publicó. Estas notas, como se podría esperar, saltan de un tema a otro y contienen cambios de notación a medida que se desarrollaba el pensamiento de Leibniz. Algunas son ideas sencillas que se le ocurrían mientras leía libros o artículos de Gregorio de St. Vincent, Fermat, Pascal, Descartes y Barrow o intentaba moldear los pensamientos de ellos en su propia forma de enfocar el cálculo. En 1714 Leibniz escribió *Historia et Origo Calculi Differentialis*, en donde proporciona una panorámica del desarrollo de su propio pensamiento. Sin embargo, esta obra fue escrita muchos años después de que hubiera realizado su trabajo y, en vista de la debilidad de la memoria humana y de la mayor percepción que había adquirido en aquella época, su historia puede no ser precisa. Como su propósito era defenderse él mismo contra una acusación de plagio, podría haber distorsionado inconscientemente su informe sobre los orígenes de sus ideas.

A pesar del confuso estado de las notas de Leibniz examinaremos algunas, porque revelan cómo uno de los mayores intelectos luchó para entender y crear. Hacia 1673 era consciente del importante problema directo e inverso de obtener tangentes a las curvas; también estaba bastante seguro de que el método inverso era equivalente al de obtener áreas y volúmenes mediante sumaciones. El desarrollo algo sistemático de sus ideas comienza con notas de 1675. Sin embargo, parece útil, para comprender su pensamiento, señalar que en su *De Arte Combinatoria* había considerado sucesiones de números, primeras diferencias, segundas diferencias y diferencias de mayor orden. Así, para la sucesión de cuadrados

$$0, 1, 4, 9, 16, 25, 36,$$

las primeras diferencias son

$$1, 3, 5, 7, 9, 11$$

y las segundas diferencias son

$$2, 2, 2, 2, 2, 2.$$

Leibniz se dio cuenta de la anulación de las segundas diferencias para la sucesión de los números naturales, la de las terceras diferencias para la sucesión de los cuadrados, y así sucesivamente. También observó, por supuesto, que si la sucesión original comienza por 0, la suma de las primeras diferencias es el último término de la sucesión.

Para relacionar estos hechos con el cálculo tuvo que pensar en la sucesión de los números como en los valores y de una función, y en la diferenciación de dos cualesquiera como la diferencia de dos valores contiguos de y. Inicialmente pensó que la x representaría el orden del término en la sucesión e y el valor de este término.

La cantidad dx, que a menudo escribe como a, es entonces 1 porque es la diferencia de los órdenes de dos términos sucesivos, y dy es la diferencia real en los valores de dos términos sucesivos. Entonces, utilizando omn. como la abreviatura del latín *omnia*, para significar suma, y utilizando l en lugar de dy, Leibniz concluye que omn. $l = y$, porque omn. l es la suma de las primeras diferencias de una sucesión cuyos términos comienzan por 0 y que, por lo tanto, proporciona el último término. Sin embargo, omn. yl presenta un

nuevo problema. Leibniz obtiene el resultado de que omn. yl es $y^2/2$ pensando en términos de la función $y = x$. Así, como muestra la figura 17.17, el área del triángulo ABC es la suma de los yl (para l pequeño) y es también $y^2/2$. Dice Leibniz: «Las líneas rectas que aumentan de la nada, multiplicada cada una por su elemento de aumento correspondiente forman un triángulo.» Estos pocos hechos ya aparecen, entre otros más complicados, en artículos de 1673.

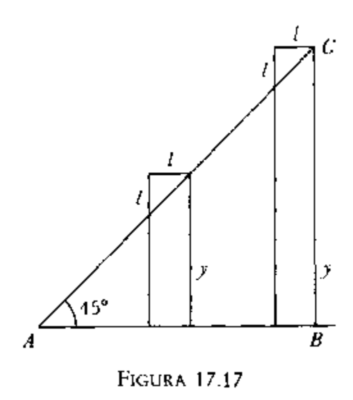


FIGURA 17.17

En el paso siguiente luchó con varias dificultades. Tuvo que efectuar la transición de una serie discreta de valores al caso en el que dy y dx son incrementos de una función arbitraria y de x. Como estaba todavía atado a las sucesiones, donde x es el orden del término, su a o dx era 1; por ello incluyó u omitió a libremente. Cuando hizo la transición al dy y dx de cualquier función, esta a no era ya 1. Sin embargo, como todavía estaba luchando con la noción de sumación, ignoró este hecho.

Así en un manuscrito del 29 de octubre de 1675, Leibniz comienza con

$$\text{omn. } yl = \overline{\text{omn.}\overline{\text{omn. } l}\frac{l}{a}}, \tag{7}$$

que se verifica porque ya y es omn. l. Aquí divide l por a para conservar las dimensiones. Leibniz dice que (7) se verifica, cualquiera que sea l. Pero, como vimos en conexión con la figura 17.17,

$$\text{omn.}\, yl = \frac{y^2}{2}\,. \tag{8}$$

Por tanto, de (7) y (8)

$$\frac{y^2}{2} = \overline{\text{omn.}\overline{\text{omn.}\, l}\,\frac{l}{a}}\,. \tag{9}$$

En nuestra notación, demostró que

$$\frac{y^2}{2} = \int\left\{\int dy\right\}\frac{dy}{dx} = \int y\,\frac{dy}{dx}\,.$$

Leibniz dice que este resultado es admirable.

Otro teorema del mismo tipo, que Leibniz obtuvo mediante razonamientos geométricos, es

$$\text{omn.}\, xl = x\ \text{omn.}\, l - \text{omn.omn.}\, l, \tag{10}$$

donde l es la diferencia en valores de dos términos sucesivos de una sucesión y x es el número del término. Para nosotros esta ecuación es

$$\int x\, dy = xy - \int y\, dx.$$

Ahora hace Leibniz x igual a l en (10), y obtiene

$$\text{omn.}\, x^2 = x\ \text{omn.}\, x - \text{omn.omn.}\, x.$$

Pero omn. x, dice, es $x^2/2$ (demostró que omn. yl es $y^2/2$). Por tanto,

$$\text{omn.}\, x^2 = x\,\frac{x^2}{2} - \text{omn.}\,\frac{x^2}{2}\,.$$

Trasponiendo el último término obtiene

$$\text{omn.}\, x^2 = \frac{x^3}{3}\,.$$

En este manuscrito del 29 de octubre de 1675, Leibniz decidió escribir ∫ en lugar de omn., de modo que

$$\int l = \text{omn.}\, l \quad \text{y} \quad \int x = \frac{x^2}{2}.$$

El símbolo ∫ es una *S* alargada para indicar «suma».

Leibniz se dio cuenta bastante pronto, probablemente al estudiar los trabajos de Barrow, de que la diferenciación y la integración como sumación deben de ser procesos inversos; así, el área, cuando se diferencia, debe proporcionar una longitud. Por ello, en el mismo manuscrito del 29 de octubre, dice Leibniz: «Dada l y su relación con x, obtener $\int l$.» A continuación, dice: «Supóngase que $\int l = ya$. Sea $l = ya/d$. [Aquí pone d en el denominador. Significaría más para nosotros si hubiera escrito $l = d(ya)$.] Entonces a medida que ∫ aumenta, d disminuye las dimensiones. Pero ∫ significa una suma y d, una diferencia. De la y dada podemos siempre obtener y/d o l, es decir, la diferencia de las y. Por lo tanto, una ecuación puede transformarse en la otra, como ocurre para la ecuación

$$\overline{\int c \int \overline{l^2}} = \frac{c \int \overline{l^3}}{3a^3},$$

de la que podemos obtener la ecuación

$$\int c \int \overline{l^2} = \frac{c \int \overline{l^3}}{3a^3 d}.»$$

En este artículo temprano, Leibniz parece estar explorando las *operaciones* de ∫ y d, y ve que son inversas. Se da cuenta al final de que ∫ no aumenta la dimensión ni d la disminuye, porque ∫ es, en realidad, una sumación de rectángulos, y por lo tanto una suma de áreas. Por lo tanto, se da cuenta de que, para volver a dy desde y, debe formar la diferencia de las y o tomar la diferencial de y. Dice entonces: «Pero ∫ significa una suma y d una diferencia.» Esto puede haber sido una inserción posterior. Así, un par de semanas más tarde, para obtener y a partir de dy, en lugar de dividir por d toma la diferencial de y, y escribe dy.

Hasta este momento Leibniz había estado pensando en los valores de *y* como en los valores de términos de una sucesión y en los de *x*, habitualmente, como en el orden de esos términos, pero ahora, en este artículo, dice: «Todos estos teoremas son ciertos para series en las que las diferencias de los términos estén, con respecto a los términos mismos, en una razón que sea menor que cualquier cantidad previamente fijada.» Es decir, que dy/y puede hacerse menor que toda cantidad cualquiera.

En un manuscrito fechado el 11 de noviembre de 1675, titulado «Ejemplos del método inverso de las tangentes», Leibniz utiliza $\int$ para la suma y x/d para la diferencia. Dice, entonces, que x/d es dx, la diferencia de dos valores de x consecutivos, pero aparentemente aquí dx es una constante e igual a la unidad.

A partir de razonamientos difícilmente inteligibles, como el anterior, Leibniz establecía el hecho de que la *integración como proceso de sumación es el inverso de la diferenciación.* Esta idea está en los trabajos de Barrow y Newton, quienes obtuvieron áreas por antidiferenciación, pero Leibniz es el primero que la expresa como una relación entre sumación y diferenciación. A pesar de su rotunda afirmación, él no tenía claro en absoluto cómo obtener un área de lo que se podría escribir vagamente como $\Sigma\, y\, dx$ —es decir, cómo obtener un área encerrada bajo una curva a partir de un conjunto de rectángulos—. Por otra parte, esta dificultad acosaba a todos los investigadores del siglo XVII. Sin poseer un concepto claro de límite, ni siquiera nociones claras sobre el área, Leibniz pensó sobre esta última a veces como una suma de rectángulos tan pequeños y tan numerosos que la diferencia entre esta suma y el área verdadera encerrada bajo la curva podía despreciarse, y otras como una suma de las ordenadas o valores de las *y*. Este último concepto de área era común, especialmente entre los indivisibilistas, quienes pensaban que la unidad de área última y el valor de *y* eran lo mismo.

Con respecto a la diferenciación, incluso después de reconocer que dy y dx pueden ser cantidades arbitrariamente pequeñas, Leibniz tenía que superar todavía la dificultad fundamental de que la razón dy/dx no es completamente la derivada en nuestro sentido. Basaba su razonamiento en el triángulo característico, que también habían utilizado Pascal y Barrow. Este triángulo (fig. 17.18) consiste en dy, dx y la cuerda PQ, que Leibniz consideró también como *la curva entre P y Q y parte de la tangente en T.* Aunque habla de este triángulo como indefinidamente pequeño, mantiene, sin embar-

go, que es semejante a un triángulo definido, el triángulo *STU* formado por la subtangente *SU*, la ordenada en *T* y la longitud de la tangente *ST*. Por lo tanto *dy* y *dx* son elementos últimos, y su cociente tiene un significado definido. De hecho, utiliza el razonamiento de que, a partir de los triángulos semejantes *PRQ* y *SUT*, $dy/dx = TU/SU$.

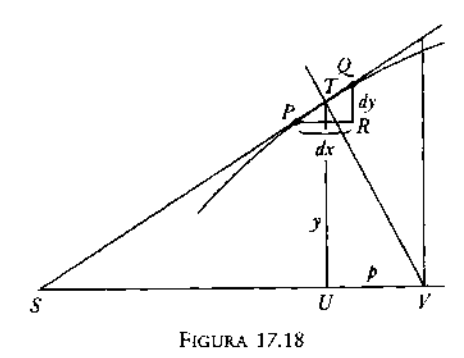


FIGURA 17.18

En el manuscrito del 11 de noviembre de 1675, Leibniz muestra cómo puede resolver un problema definido. Busca la curva cuya subnormal es inversamente proporcional a la ordenada. En la figura 17.18, la normal es *TV* y la subnormal *p* es *UV*. De la semejanza de los triángulos *PRQ* y *TUV* obtiene

$$\frac{dy}{dx} = \frac{p}{y}$$

o bien

$$p\,dx = y\,dy.$$

Pero la curva tiene la propiedad dada de que

$$p = \frac{b}{y},$$

donde *b* es la constante de proporcionalidad. Por tanto,

$$dx = \frac{y^2}{b}\,dy.$$

Entonces

$$\int dx = \int \frac{y^2}{b}\,dy$$

o bien

$$x = \frac{y^3}{3b}\,.$$

Leibniz también resolvió otros problemas inversos de tangentes.

En un artículo del 26 de junio de 1676, se da cuenta de que el mejor método para obtener las tangentes es hallar dy/dx, donde dy y dx son diferencias y dy/dx es el cociente. Ignora $dx.dx$ y potencias de orden superior de dx.

En noviembre de 1676 es capaz de dar las reglas generales $dx^n = nx^{n-1}$ para n entero y fraccionario y $\int x^n = x^{n+1}/(n+1)$, y dice: «El razonamiento es general, y no depende de cuáles puedan ser las progresiones de las x.» Aquí x todavía significa el orden de los términos de una sucesión. En este manuscrito también dice que para diferenciar $\sqrt{a + bz + cz^2}$, se hace $a + bz + cz^2 = x$, se diferencia $\sqrt{x}$, y se multiplica por dx/dz. Esta es la regla de la cadena.

El 11 de julio de 1677, Leibniz podía dar las reglas correctas para la diferencial de la suma, diferencia, producto y cociente de dos funciones y para las potencias y raíces, pero sin demostraciones. En el manuscrito del 11 de noviembre de 1675 se había peleado con $d\,(u\,v)$ y $d\,(u/v)$, y pensó que $d\,(u\,v) = du\,dv$.

En 1680, dx se había convertido en la diferencia de las abscisas y dy en la diferencia de las ordenadas. Dice: «... ahora, estas dx y dy se tomarán como infinitamente pequeñas, o bien se entiende que los dos puntos de la curva están separados una distancia que es menor que cualquier longitud dada...». Llama a dy el «incremento momentáneo» en y cuando la ordenada se mueve a lo largo del eje de las x. Pero PQ, en la figura 17.18, se considera todavía como parte de una línea recta. Es «un elemento de la curva o un lado de un polígono de infinitos ángulos que se apoya en la curva...». Con-

tinúa utilizando la forma diferencial habitual. Por tanto, sí $y = a^2/x$, entonces

$$dy = -\frac{a^2}{x^2}\,dx.$$

Dice también que las diferencias son lo opuesto de las sumas. Por lo tanto, para obtener el área encerrada bajo una curva (fig. 17.19), toma la suma de los rectángulos y dice que se puede despreciar todos los restantes «triángulos, puesto que son infinitamente pequeños comparados con los rectángulos (anteriores)... por tanto en mi cálculo representado el área de la figura mediante $\int y\,dx$...». También proporciona, para el elemento de arco

$$ds = \sqrt{dx^2 + dy^2}\,;$$

y para el volumen de un sólido de revolución obtenido por el giro de una curva alrededor del eje x

$$V = \pi \int y^2\,dx.$$

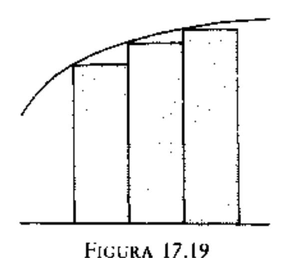


FIGURA 17.19

A pesar de algunas afirmaciones previas de que dx y dy son diferencias pequeñas, Leibniz todavía habla de sucesiones. Dice que «Las diferencias y sumas son inversas unas de otras, es decir, la suma de las diferencias de una serie [sucesión] es un término de la serie, y la diferencia de las sumas de una serie es un término de la serie, y enumero lo primero así, $\int dx = x$, y lo último así, $d \int x = dx$». De hecho, en un manuscrito escrito después de 1684, Leibniz dice que

su método de los infinitesimales ha llegado a ser ampliamente conocido como el cálculo de diferencias.

La primera publicación de Leibniz sobre el cálculo está en el *Acta Eruditorum* de 1684 [12]. En este artículo, el significado de dy y de dx no está todavía claro. Define, en este artículo, dx como una cantidad arbitraria, y dy como (ver la fig. 17.18).

$$dy : dx = y : \text{subtangente}.$$

Esta definición de dy presupone alguna expresión para la subtangente; por tanto la definición no es completa. Además, la definición de Leibniz de la tangente como una recta que une dos puntos infinitamente próximos no es satisfactoria.

También da en este artículo las reglas que había obtenido en 1677 para la diferencial de la suma, producto y cociente de dos funciones, y la regla para obtener $d\,(x^n)$. En este último caso, esboza la demostración para n entero positivo, pero dice que la regla es verdadera para todo n; para las otras reglas no da demostraciones. Hace aplicaciones a la obtención de tangentes, máximos y mínimos y puntos de inflexión. Este artículo, de seis páginas de largo, es tan poco claro que los hermanos Bernoulli lo llamaron «un enigma más que una explicación» [13].

En su artículo de 1686 [14] Leibniz da

$$y = \sqrt{2x - x^2} + \int \frac{dx}{\sqrt{2x - x^2}}$$

como la ecuación de la cicloide. Desea demostrar que, mediante sus métodos y notación, algunas curvas pueden expresarse como ecuaciones que no se pueden obtener de otra manera. Reafirma esto mismo en su *Historia,* donde dice que sus dx, ddx (segunda diferencia) y las sumas que son las inversas de estas diferencias pueden aplicarse a todas las funciones de x, sin exceptuar a las curvas mecánicas de Vieta y Descartes, de las cuáles éste había dicho que no tenían ecuación. Leibniz dice también que puede incluir curvas que Newton no podría manejar ni siquiera con su método de las series.

---

[12] *Acta Erud.*, 3, 1684, 467-73 = *Math. Schriften*, 5, 220-26.

[13] Leibniz: *Math. Schriften*, 3, Parte I, 5.

[14] *Acta Erud.*, 5, 1686, 292-300 = *Math. Schriften*, 5, 226-233.

En el artículo de 1686 así como en artículos posteriores [15], Leibniz dio las diferenciales de las funciones exponencial y logarítmica, e identificó las funciones exponenciales como una clase. También trató la curvatura, el círculo osculador y la teoría de las envolventes (ver el cap. 23). En una carta de Jean Bernoulli en 1697, diferenciaba bajo el signo integral con respecto a un parámetro. También tuvo la idea de que muchas integrales indefinidas podían ser calculadas reduciéndolas a formas conocidas, y habla de preparar tablas para tales reducciones —en otras palabras, una tabla de integrales—. Intentó definir las diferenciales de orden superior, tales como *ddy* (d^2y) y *dddy* (d^3y), pero las definiciones no fueron satisfactorias. Aunque sin éxito, intentó obtener un significado para d^α y donde α es un número real cualquiera.

Con respecto a la notación, Leibniz trabajó esmeradamente en busca de la más adecuada. Sus *dx, dy* y *dy/dx* se utilizan todavía. Introdujo la notación log *x*, d^n para la diferencial *n-ésima,* e incluso d^{-1} y d^{-n} para $\int$ y para la iteración *n-ésima* de la sumación, respectivamente.

En general, el trabajo de Leibniz, aunque rico en sugestiones y profundo, fue tan incompleto y fragmentario que resultaba difícilmente inteligible. Afortunadamente los hermanos Bernoulli, Jacques y Jean, se impresionaron y estimularon inmensamente con las ideas de Leibniz, pulieron sus esquemáticos trabajos y aportaron una cantidad inmensa de nuevos desarrollos que trataremos más adelante. Leibniz reconoció que el cálculo era tanto de ellos como suyo.

5. Una comparación de las obras de Newton y Leibniz

Tanto a Newton como a Leibniz se les debe reconocer que vieran el cálculo como un método nuevo y general, aplicable a muchos tipos de funciones. Después de su contribución, el cálculo dejó de ser un apéndice y una extensión de la geometría griega para convertirse en una ciencia independiente capaz de manejar una cantidad de problemas ampliamente extendida.

También ambos aritmetizaron el cálculo, es decir, construyeron en él con conceptos algebraicos. La notación y técnicas algebraicas

---

[15] *Acta Erud.*, 1692, 168-171 = *Math. Schriften*, 5, 266-269; *Acta Erud.*, 1694 = *Math. Schriften*, 5, 301-306.

utilizadas por Newton y Leibniz no sólo les proporcionaron un instrumento más efectivo que la geometría, sino que también permitieron tratar con la misma técnica muchos problemas geométricos y físicos diferentes. Un cambio importante que se produjo del comienzo al fin del siglo XVII fue la algebrización del cálculo. Este fenómeno es comparable a lo que Vieta había hecho en la teoría de ecuaciones y Descartes y Fermat en geometría.

La tercera contribución vital que comparten Newton y Leibniz es la reducción a la antidiferenciación del área, volumen y otros problemas que habían sido tratados anteriormente como sumaciones. Así, los cuatro problemas principales —cambio relativo, tangentes, máximos y mínimos y sumación— quedaron reducidos todos a diferenciación y antidiferenciación.

La distinción principal entre el trabajo de los dos es que Newton utilizó los incrementos infinitamente pequeños en x y en y como medio para determinar la fluxión o derivada. Era esencialmente el límite del cociente de los incrementos, cuando éstos se hacían cada vez más pequeños. Por otra parte, Leibniz trató directamente con los incrementos infinitamente pequeños en x y en y, es decir, con diferenciales, y determinó las relaciones entre ellos. Esta diferencia refleja la orientación física de Newton, en la que un concepto como el de velocidad es central, y la preocupación filosófica de Leibniz por las partículas últimas de la materia, que llamó mónadas. Como consecuencia, Newton resolvió los problemas de áreas y volúmenes pensando enteramente en términos de cambio relativo. Para él, la diferenciación era básica; este proceso y su inverso resolvían todos los problemas del cálculo y, de hecho, el uso de la sumación para obtener un área, un volumen o un centro de gravedad aparece raramente en sus trabajos. Leibniz, en cambio, pensaba primero en términos de sumación aunque, por supuesto, estas sumaciones se calcularan mediante antidiferenciación.

Una tercera distinción entre las obras de los dos está en la libre utilización, por parte de Newton, de las series para representar funciones; Leibniz prefería la forma cerrada. En una carta a Leibniz en 1676, Newton hacía hincapié en el uso de las series, incluso para resolver ecuaciones diferenciales sencillas. Aunque Leibniz utilizaba series infinitas, replicó que el objetivo real era el de obtener los resultados en términos finitos, utilizando las funciones trigonométricas y logarítmicas cuando no sirvieran las funciones algebraicas. Recordó a Newton la afirmación de James Gregory de que la recti-

ficación de la elipse y de la hipérbola no podía reducirse a las funciones circulares y logarítmica, y retó a Newton para que determinara, mediante el uso de las series, si Gregory estaba en lo cierto o no. Newton respondió que mediante la utilización de las series podía decidir si algunas integraciones podían realizarse en términos finitos, pero no dio ningún criterio. De nuevo, en una carta a Jean Bernoulli fechada en 1712, Leibniz ponía objeciones al desarrollo en serie de funciones, y afirmaba que el cálculo debía ocuparse de reducir sus resultados a cuadraturas (integraciones) y, donde fuera necesario, cuadraturas que incluyeran funciones trascendentes.

Hay diferencias en sus maneras de trabajar. Newton era empírico, concreto y circunspecto, mientras que Leibniz era especulativo, dado a las generalizaciones y osado. Leibniz estaba más preocupado por las fórmulas operacionales para elaborar un cálculo en sentido amplio; por ejemplo, reglas para la diferencial de un producto o cociente de funciones, su regla para $d^n (u\, v)$ (donde u y v son funciones de x) y una tabla de integrales. Fue Leibniz quien estableció los cánones del cálculo, el sistema de reglas y fórmulas. Newton no se molestó en formular reglas, aun cuando pudo haber generalizado fácilmente sus resultados concretos. Sabía que si $z = uv$, entonces $\dot{z} = u\dot{v} + v\dot{u}$, pero no resaltó este resultado general. Aunque Newton inició muchos métodos, no hizo hincapié en ellos. Sus grandiosas aplicaciones del cálculo no sólo demostraron su valor sino que estimularon y casi determinaron enteramente la dirección del análisis del siglo XVIII, en mucha mayor medida que el trabajo de Leibniz. Newton y Leibniz difirieron también en su preocupación por la notación. Newton no daba importancia a este asunto, mientras que Leibniz dedicaba días enteros a elegir una notación sugestiva.

6. La controversia sobre la prioridad

Nada del trabajo de Newton sobre el cálculo fue publicado antes de 1687, aunque había comunicado diversos resultados a amigos entre los años 1665 y 1687. En particular, había enviado su tratado *De Analysi* en 1669 a Barrow, quien lo envió a John Collins. Leibniz visitó París en 1672 y Londres en 1673 y se comunicó con gente que conocía el trabajo de Newton. Sin embargo, no publicó nada sobre el cálculo hasta 1684. Por tanto, se suscitó la cuestión de si Leibniz había conocido los detalles de lo que había hecho Newton,

y Leibniz fue acusado de plagio. Sin embargo, investigaciones realizadas mucho después de la muerte de los dos mostraron que Leibniz descubrió independientemente ideas importantes sobre el cálculo, aunque Newton realizó la mayor parte de su trabajo antes que Leibniz. Ambos deben mucho a Barrow, aunque éste utilizó casi exclusivamente métodos geométricos. El significado de la controversia no está en la cuestión de quién fue el triunfador sino más bien en el hecho de que los matemáticos tomaron partido. Los matemáticos continentales, los hermanos Bernoulli en particular, se pusieron del lado de Leibniz, mientras que los matemáticos ingleses defendieron a Newton. Los dos grupos se enemistaron, incluso agriamente, entre sí; Jean Bernoulli llegó a ridiculizar y lanzar invectivas contra de los ingleses.

El resultado fue que los matemáticos ingleses y continentales cesaron de intercambiar ideas. Como el trabajo principal de Newton y primera publicación sobre el cálculo, los *Principia*, utilizaba métodos geométricos, los ingleses continuaron utilizando principalmente la geometría durante cerca de cien años después de su muerte. Los continentales adoptaron los métodos analíticos de Leibniz y los ampliaron y mejoraron. Estos demostraron ser bastante más efectivos; por ello, no sólo los matemáticos ingleses se quedaron más atrás, sino que las matemáticas se vieron privadas de contribuciones que podrían haber realizado las mentes más capaces.

7. Algunas adiciones inmediatas al cálculo

El cálculo es, por supuesto, el comienzo de la muy importante parte de las matemáticas que se conoce como análisis. Seguiremos los importantes desarrollos de este campo en los capítulos siguientes; señalaremos aquí, no obstante, algunas adiciones que se hicieron inmediatamente después del trabajo básico de Newton y Leibniz.

En su *Arithmetica Universalis* (1707) Newton estableció un teorema sobre la cota superior de las raíces reales de ecuaciones polinómicas. El teorema dice: un número a es una cota superior de las raíces reales de $f(x) = 0$ si, cuando se sustituye x por a se tiene que $f(x)$ y todas sus derivadas mantienen el mismo signo.

En su *De Analysi* y su *Method of Fluxions*, dio un método general para aproximar las raíces de $f(x) = 0$, que fue publicado en el *Algebra* de Wallis de 1685. En su tratado *Analysis Aequationum*

Universalis (1690), Joseph Raphson (1648-1715) mejoró este método; aunque lo aplicó sólo a polinomios, es mucho más ampliamente utilizable. Esta modificación es conocida ahora como el método de Newton o el método de Newton-Raphson. Consiste en tomar una aproximación a, en primer lugar, y calcular a continuación $a - f(a)/f'(a)$. Se llama a esto b y se calcula $b - f(b)/f'(b)$. Se llama c a este resultado y así sucesivamente. Los números $a, b, c, \ldots$ son aproximaciones sucesivas de la raíz. (La notación es moderna.) En realidad, el método no proporciona necesariamente mejores aproximaciones de la raíz J. Raymond Mourraille mostró en 1768 que a debe escogerse de manera que la curva $y = f(x)$ sea convexa hacia el eje de las x en el intervalo entre a y la raíz. Bastante más tarde Fourier descubrió este hecho independientemente.

En su *Démostration d'une méthode pour résoudre les égalitéz de tous les dégrez* (1691), Michel Rolle (1652-1719) incluyó el famoso teorema que lleva su nombre, es decir, el que establece que si una función es cero en dos valores de x, a y b, entonces la derivada es cero en algún valor de x entre a y b. Rolle enunció el teorema pero no lo demostró.

Después de Newton y Leibniz, los dos fundadores más importantes del cálculo fueron los hermanos Bernoulli, Jacques y Jean. Jacques Bernoulli (1655-1705) fue autodidacta en matemáticas y por eso maduró lentamente en ellas. A instancias de su padre estudió para el sacerdocio, pero finalmente se orientó hacia las matemáticas y en 1686 se hizo profesor en la Universidad de Basilea. Sus intereses principales desde entonces fueron las matemáticas y la astronomía. Cuando, a finales de los 1670, comenzó a trabajar en problemas matemáticos, los trabajos de Newton y Leibniz eran todavía desconocidos para él. También aprendió de *La Géométrie* de Descartes, de la *Arithmetica Infinitorum* de Wallis y de las *Geometrical Lectures* de Barrow. Aunque tomó muchas cosas de Barrow, las puso en forma analítica. Se familiarizó gradualmente con el trabajo de Leibniz, pero como había aparecido tan poco impreso, mucho de lo que hizo Jacques coincidió con los resultados de Leibniz. En realidad, como ocurrió con otros matemáticos de su tiempo, no comprendió completamente el trabajo de Leibniz.

La actividad de Jacques está íntimamente relacionada con la de su hermano menor Jean (1667-1748). A Jean le dedicó su padre a los negocios, pero se orientó hacia la medicina, mientras aprendía matemáticas de su hermano. Se hizo profesor de matemáticas en

Groningen, Holanda, y después sucedió a su hermano en Basilea.

Tanto Jacques como Jean mantuvieron correspondencia constante con Leibniz, con Huygens, con otros matemáticos y entre ellos mismos. Todos trabajaban en muchos problemas comunes, sugeridos en cartas o propuestos como desafíos. Como también los resultados, en aquella época, se comunicaban a menudo por carta, con o sin publicación ulterior, el asunto de la prioridad es complicado. A veces se reclamaba el mérito de un resultado, que había sido anunciado, pero del que no se daba la demostración en ese momento. La situación se complicaba más por la peculiares relaciones que se desarrollaron. Jean estaba extremadamente ansioso por conseguir la fama y comenzó a competir con su hermano; pronto comenzaron a desafiarse con problemas. Jean no dudó en utilizar medios poco escrupulosos para aparecer como el descubridor de resultados que había obtenido de otros, incluyendo su heramno. Jacques era muy sensible y reaccionó de la misma manera. Los dos publicaron artículos que debían mucho al otro, sin reconocer el origen de las ideas. De hecho, Jean se convirtió en un crítico acérrimo de su hermano, y Leibniz trató de mediar entre ambos. Aunque Jacques había dicho anteriormente, mientras alababa a Barrow, que el trabajo de Leibniz no debía ser menospreciado, se hizo cada vez más receloso con respecto a Leibniz. Además le molestaba la superior percepción de Leibniz y pensaba que éste era muy arrogante al señalar que había hecho cosas que él pensaba que eran suyas. Llegó a convencerse de que Leibniz sólo quería minimizar su trabajo y que estaba favoreciendo a Jean en las disputas entre los hermanos. Cuando Nicolás Fatio de Duillier (1664-1753) le atribuyó a Newton el mérito de la creación del cálculo y se enzarzó en una controversia con Leibniz, Jacques escribió cartas a Fatio oponiéndose a Leibniz.

Por lo que se refiere al trabajo de los Bernoulli en el cálculo, se puede decir que ellos también trataron problemas tales como la obtención de curvaturas de curvas, evolutas (envolventes de las normales a una curva), puntos de inflexión, rectificación de curvas y otros temas básicos del cálculo. Los resultados de Newton y Leibniz fueron extendidos a espirales de distintos tipos, a la catenaria y a la tractriz, que fue definida como la curva (fig. 17.20) para la cual el cociente entre PT y OT es una constante. Jacques también escribió cinco artículos importantes sobre series (cap. 20, sec. 4), que ampliaron la utilización de Newton de las series para integrar funciones algebraicas complicadas y trascendentes. En 1691, tanto Jacques como

Jean dieron la fórmula del radio de curvatura de una curva. Jacques llamó a este resultado el «teorema áureo» y lo escribió como

$$z = dx\, ds : ddy = dy\, ds : ddx$$

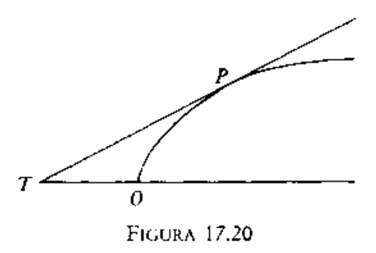


FIGURA 17.20

donde z es el radio de curvatura. Si dividimos el numerador y el denominador de cada uno de los cocientes por ds^2 obtenemos

$$z = \frac{dx/ds}{d^2y/ds^2} = \frac{dy/ds}{d^2x/ds^2},$$

que son formas más familiares. Jacques también dio el resultado en coordenadas polares.

Jean produjo un teorema, ahora famoso, para obtener el límite al que se aproxima una fracción cuyos numerador y denominador se acercan a cero. Este teorema fue incluido por Guillaume F. A. l'Hôpital (1661-1704), un discípulo de Jean, en un influyente libro sobre el cálculo, el *Analyse des infiniment petits* (1696), y se conoce ahora como regla de l'Hôpital.

8. La solidez del cálculo

Desde el mismo momento de la presentación de los nuevos métodos para obtener cambios relativos, tangentes, máximos y mínimos, etc., se atacó a las demostraciones como poco sólidas. El uso de los elementos últimos indivisibles de Cavalieri y sus razonamientos extrañaron a quienes todavía respetaban el rigor lógico. A sus críticas, Cavalieri respondía que los geómetras contemporáneos habían sido más libres con respecto a la lógica que él —por ejemplo, Kepler, en su *Stereometria Doliorum*—. Estos geómetras, continua-

ba, se habían limitado, en su cálculo de áreas, a imitar el método de Arquímedes de sumar líneas, pero no habían proporcionado las demostraciones rigurosas que los prestigiosos griegos habían utilizado para hacer riguroso su trabajo. Estaban satisfechos con sus cálculos sólo con tal de que los resultados fueran útiles. Cavalieri se sintió justificado al adoptar el mismo punto de vista. Decía que sus procedimientos podían conducir a nuevos inventos y que su método no obligaba en absoluto a considerar una estructura geométrica como compuesta por un número infinito de secciones; no tenía otro objetivo que el de establecer razones correctas entre áreas o volúmenes. Esas razones conservarían su sentido y valor cualquiera que fuera la opinión que se pudiera tener sobre la composición de un continuo. En cualquier caso, decía Cavalieri, «el rigor es la preocupación de la filosofía y no de la geometría».

Fermat, Pascal y Barrow se dieron cuenta de la imprecisión de sus trabajos sobre sumación, pero creían que se podían hacer demostraciones precisas a la manera de Arquímedes. Pascal, en las *Cartas de Dettonville* (1659), afirmaba que la geometría infinitesimal y la geometría clásica griega estaban en buen acuerdo. Concluía: «lo que se demuestra mediante las reglas verdaderas de los indivisibles podría también demostrarse con el rigor y en la forma de los antiguos». Además, decía que el método de los indivisibles debe ser aceptado por cualquier matemático que pretenda contarse entre los geómetras. Difiere del método de los antiguos sólo en el lenguaje. Sin embargo, también Pascal tenía sentimientos ambivalentes acerca del rigor. A veces opinaba que el corazón interviene para asegurarnos la corrección de los pasos matemáticos. El «discernimiento» adecuado, más que la lógica geométrica, es lo que se necesita para realizar un trabajo correcto, así como la valoración religiosa de que la gracia está por encima de la razón. Las paradojas de la geometría, tal como se utilizan en el cálculo, son como los aparentes absurdos del cristianismo, y el indivisble en geometría está en la misma relación con lo finito que la justicia del hombre con la de Dios.

La resistencia que ofrecieron Cavalieri y Pascal se refería a la sumación de cantidades infinitamente pequeñas. Como con la derivada, los primeros investigadores como Fermat y Roberval pensaron que se trataba de un proceso algebraico sencillo que tenía una interpretación geométrica clara y que, por lo tanto, podía justificarse mediante razonamientos geométricos. En realidad Fermat tuvo mucho cuidado de no enunciar ningún teorema general cuando adelan-

taba alguna idea que no podía justificar por el método exhaustivo. Barrow razonaba sólo geométricamente y, a pesar de sus ataques a los algebristas por su falta de rigor, él era menos escrupuloso acerca de la solidez de sus razonamientos geométricos.

Ni Newton ni Leibniz entendieron claramente, ni definieron rigurosamente, sus conceptos fundamentales. Hemos observado ya que ambos vacilaron en sus definiciones de la derivada y de las diferenciales. Newton no creyó en realidad que había partido de la geometría griega. Aunque utilizó el álgebra y la geometría de coordenadas, de las que no gustaba mucho, pensó que los métodos subyacentes no eran más que extensiones naturales de la geometría pura. Leibniz, sin embargo, era un hombre de visión que pensó en términos amplios, como Descartes. Vio las implicaciones a largo plazo de las nuevas ideas, y no dudó en declarar que estaba naciendo una nueva ciencia. Por ello no estaba demasiado preocupado por la falta de rigor en el cálculo.

En respuesta a la crítica a sus ideas, Leibniz escribió varias réplicas, insatisfactorias. En una carta de Wallis, de 30 de marzo de 1690 [16], dice que

> Es útil considerar cantidades infinitamente pequeñas tales que, cuando se busca su cociente, pueden no considerarse cero, pero que pueden despreciarse cuando aparecen con cantidades incomparablemente más grandes. Por lo tanto, si tenemos $x + dx$, dx puede despreciarse. Pero es diferente si buscamos la diferencia entre $x + dx$ y x. De la misma manera, no podemos tener a la vez $x\,dx$ y $dx\,dx$. Por lo tanto, si tenemos que diferenciar xy, escribimos $(x + dx)\,(x + dx) - xy = x\,dy + y\,dx + dx\,dy$. Pero $dx\,dy$ puede despreciarse como incomparablemente menor que $x\,dy + y\,dx$. Así, en cualquier caso particular, el error es menor que cualquier cantidad finita.

Por lo que se refiere a los significados últimos de dy, dx y dy/dx, Leibniz fue siempre impreciso. Hablaba de dx como de la diferencia de los valores de x entre dos puntos infinitamente próximos y de la tangente como de la recta que une tales puntos. Despreció diferenciales de orden superior sin ninguna justificación, aunque distinguía entre los distintos órdenes. Los infinitamente pequeños dx y dy se describían a veces como tendiendo a cero o como cantidades incipientes, como opuestas a cantidades ya formadas. Estas cantidades indefinidamente pequeñas no eran cero, pero eran más pequeñas que

[16] Leibniz: *Math. Schriften,* 4, 63.

cualquier cantidad finita. A veces recurría a la geometría para decir que un diferencial de orden superior es a uno de orden inferior como un punto es a una línea [17] o que *dx* es a *x* como un punto a la Tierra o como el radio de ésta al de los cielos. Pensó en el cociente de dos infinitesimales como en el de inasignables o de cantidades indefinidamente pequeñas, pero de manera que pudiera, sin embargo, ser expresado en términos de cantidades definidas tales como el cociente de la ordenada a la subtangente.

Una tormenta de ataques y refutaciones se inició en libros de 1694 y 1695 por el físico y geómetra holandés Bernard Nieuwentijdt (1654-1718), Aunque admitía que los nuevos métodos, en general, conducían a resultados correctos, criticaba la oscuridad de los mismos y señalaba que a veces conducían a absurdos. Se quejaba de que no podía entender cómo las cantidades infinitamente pequeñas se diferenciaban de cero y preguntaba cómo una suma de infinitesimales podía ser finita. También dudaba del significado y de la existencia de diferenciales de orden superior y de poder despreciar cantidades infinitamente pequeñas en partes de los razonamientos.

Leibniz, en un borrador de una respuesta de Nieuwentijdt, escrito probablemente en 1695, y en un artículo en el *Acta Eruditorum* de 1695 [18], da varias respuestas. Habla de críticas «sobreprecisas» y dice que un exceso de escrúpulos no debería inducirnos a rechazar los frutos de la invención. Dice a continuación que sus métodos difieren de los de Arquímedes sólo en las expresiones utilizadas, pero que los suyos están mejor adaptados al arte del descubrimiento. Las palabras «infinito» e «infinitesimal» significan meramente cantidades que se pueden tomar tan grandes o tan pequeñas como se desee para mostrar que el error en que se incurre es menor que cualquier número que pueda fijarse de antemano —en otras palabras, que no hay error—. Se puede utilizar estos entes últimos —esto es, cantidades infinitas e infinitamente pequeñas— como un instrumento, en la misma forma en que los algebristas utilizaban las raíces imaginarias con gran provecho.

Los razonamientos de Leibniz hasta entonces consistían en que su cálculo utilizaba sólo los conceptos matemáticos ordinarios. Pero como no pudo satisfacer adecuadamente a sus críticos, enunció un principio filosófico conocido como la ley de continuidad, que era

---

[17] *Math. Schriften*, 5, 322 y sig.

[18] *Acta Erud.*, 1695, 310-316 = *Math. Schriften*, 5, 320-328.

prácticamente el mismo ya establecido por Kepler. En 1687, en una carta a Pierre Bayle [19], Leibniz expresó este principio de la manera siguiente: «En cualquier supuesta transición que acaba en un término, es válido elaborar un razonamiento general en el que el término final esté incluido.» Para apoyar este principio proporciona, en un manuscrito no publicado de, aproximadamente, 1695, el ejemplo de incluir en un razonamiento elipses y parábolas, aunque la parábola es un caso límite de la elipse cuando uno de los focos se desplaza hasta el infinito. Aplica entonces el principio al cálculo de $dy{:}dx$ para la parábola $y = x^2/a$. Después de obtener

$$dy : dx = (2x + dx) : a,$$

dice: «Ahora, como por nuestro postulado es posible incluir en el razonamiento general también el caso en el que (fig. 17.21) la ordenada x_2y_2 se desplaza acercándose cada vez más a la x_1y_1 hasta que llega a coincidir con ella, es evidente que en este caso dx se hace cero y debe ser despreciado...» Leibniz no dice qué significado debe dársele al dx que aparece en el primer miembro de la ecuación.

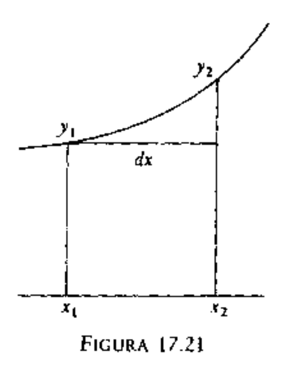


FIGURA 17.21

Por supuesto, dice, que cosas que son absolutamente iguales tienen una diferencia nula; por lo tanto, una parábola no es una elipse.

Todavía puede imaginarse un estado de transición o de evanescencia en el que aún no se haya llegado exactamente a la igualdad o al reposo... pero en el que se esté pasando a un estado tal que la diferencia sea menor que

[19] *Math. Schriften,* 5, 385.

cualquier cantidad asignable; en ese estado puede todavía quedar una diferencia [con respecto a la igualdad], alguna velocidad, algún ángulo (distinto de cero), pero en todo caso algo infinitamente pequeño...

Por el momento, tanto el que un tal estado de transición entre la desigualdad y la igualdad... pueda sostenerse en un sentido riguroso o metafísico, o que extensiones infinitas cada vez más grandes o infinitamente pequeñas cada vez menores sean consideraciones legítimas, es un asunto que considero posiblemente cuestionable...

Será suficiente si, cuando hablamos de cantidades infinitamente grandes (o, más estrictamente, ilimitadas) o infinitamente pequeñas (es decir, las menores a las que alcance nuestro conocimiento), se entiende que significamos cantidades que son indefinidamente grandes o indefinidamente pequeñas, es decir, tan grandes o tan pequeñas como se quiera, de modo que el error que se pueda asignar previamente sea menor que una cantidad establecida de antemano.

En estos supuestos, todas las reglas de nuestros algoritmos, en la forma en que están establecidas en el *Acta Eruditorum* de octubre de 1684, pueden probarse sin demasiada dificultad.

Leibniz vuelve entonces sobre esas reglas. Introduce las cantidades *(d)y* y *(d)x* y lleva a cabo el procedimiento habitual de diferenciación con ellas. Llama a éstas cantidades asignables o definidas no evanescentes. Después de obtener el resultado final, dice, podemos sustituir *(d)y* y *(d)x* por las cantidades evanescentes o no asignables *dy* y *dx*, haciendo «la suposición de que el cociente de las cantidades evanescentes *dy* y *dx* es igual al cociente de *(d)y* y *(d)x*, porque esta suposición puede siempre ser reducida a una verdad indudable».

El principio de continuidad de Leibniz no es hoy, ciertamente, un axioma matemático, pero él le dio importancia, y resultó importante más tarde. Muchos de sus razonamientos están de acuerdo con este principio. Por ejemplo, en una carta a Wallis [20], Leibniz defendía su utilización del triángulo característico como una forma sin magnitud, la forma que quedaba después de que las magnitudes habían sido reducidas a cero, y preguntaba, desafiante: «¿Quién no admite una forma sin magnitud?» De la misma manera, en una carta a Guido Grandi [21], dice que lo infinitamente pequeño no es un cero simple y absoluto sino un cero relativo, es decir, una cantidad evanescente que mantiene todavía el carácter de la que está desapare-

---

[20] *Math. Schriften*, 4, 54.
[21] *Math. Schriften*, 4, 218.

ciendo. Sin embargo, dice también Leibniz, en otras ocasiones, que no cree en magnitudes verdaderamente infinitas o verdaderamente infinitesimales.

Leibniz, menos preocupado por la justificación última de sus procedimientos que Newton, sentía que aquélla se apoyaba en la efectividad de éstos. Acentuaba el valor algorítmico o de procedimiento de lo que había creado. De alguna manera tenía confianza en que si formulaba claramente las reglas de operación, y si éstas se aplicaban adecuadamente, se obtendrían resultados razonables y correctos, aunque pudieran ser dudosos los significados de los símbolos relacionados.

Parece evidente que ni Newton ni Leibniz lograron clarificar, y mucho menos precisar, los conceptos básicos del cálculo: la derivada y la integral. No siendo capaces de dominarlos adecuadamente, confiaron en la coherencia de los resultados y la fecundidad de los métodos para seguir adelante sin rigor.

Varios ejemplos pueden ilustrar esta falta de claridad, incluso entre los más grandes sucesores inmediatos de Newton y Leibniz. Jean Bernoulli escribió su primer texto sobre el cálculo en 1691 y 1692. La parte de cálculo integral fue publicada en 1742 [22]; La parte de cálculo diferencial, *Die Differentialrechnung*, no fue publicada hasta 1924. Sin embargo, el Marqués de l'Hôpital publicó una versión francesa ligeramente modificada (a la que ya se ha hecho referencia) bajo su propio nombre en 1696. Bernoulli comienza el *Differentialrechnung* con tres postulados. El primero dice así: «Una cantidad que decrece o aumenta en una cantidad infinitamente pequeña ni se incrementa ni disminuye.» Su segundo postulado es: «Toda línea curva consta de infinitas líneas rectas, que son infinitamente pequeñas.» En su razonamiento sigue a Leibniz y utiliza los infinitesimales. Así, para obtener dy a partir de $y = x^2$, utiliza e en lugar de dx y obtiene $(x + e)^2 - x^2$, o $2xe + e^2$, y entonces suprime e^2. Como Leibniz, utilizó vagas analogías para explicar lo que eran las diferenciales. Así, dice, las cantidades infinitamente grandes son como distancias astronómicas y las infinitamente pequeñas son como animalillos descubiertos en el microscopio. En 1698 estableció que los infinitesimales deben existir [23]. Sólo se tiene que considerar la serie infinita 1, 1/2, 1/4, ... Si se toman 10 términos, entonces 1/10

[22] *Opera Omnia*, 3, 385-558.

[23] Leibniz: *Math. Schriften*, 3, Parte 2, 563 y sigs.

existe; si se toman 100 términos, 1/100 existe. El infinitesimal corresponde al número infinito de términos.

Unos pocos, Wallis y Jean Bernoulli entre ellos, intentaron definir el infinitesimal como el recíproco de ∞, porque éste era un número definido para ellos. Aún otros actuaban como si lo que era incomprensible no necesitara explicación adicional. Para la mayor parte de los investigadores del siglo XVIII, el rigor no era motivo de preocupación. Lo que decían, a menudo, que podía hacerse riguroso mediante el método de Arquímedes, no podría haberlo rigorizado un Arquímedes; esto es particularmente cierto por lo que se refiere al trabajo sobre la diferenciación, que no tenía paralelo en la matemática griega.

En realidad, el nuevo cálculo estaba introduciendo conceptos y métodos que inauguraban una separación radical del trabajo previo. Con el trabajo de Newton y Leibniz, el cálculo se convirtió en una disciplina totalmente nueva que requería sus propios cimientos. Aunque no eran totalmente conscientes de ello, los matemáticos habían vuelto la espalda al pasado.

Gérmenes de los nuevos conceptos correctos pueden encontrarse incluso en la literatura del siglo XVII. Wallis, en la *Aritmetica Infinitorum,* avanzó el concepto aritmético de límite de una función como un número al que se aproxima la función de modo que la diferencia entre este número y la función puede hacerse menor que cualquier cantidad fijada de antemano y que se anularía cuando el proceso se continuara hasta el infinito. La forma de decirlo es vaga, pero contiene la idea correcta.

James Gregory, en su *Vera Circuli et Hiperbolae Quadratura* (1667), señalaba explícitamente que los métodos utilizados para obtener áreas, volúmenes y longitudes de curvas incluían un nuevo proceso, el proceso de paso al límite. Además, añadía, esta operación era distinta de las cinco operaciones algebraicas de adición, sustracción, multiplicación, división, y extracción de raíces. Dio forma algebraica al método exhaustivo y se dio cuenta de que las aproximaciones sucesivas obtenidas utilizando figuras rectilíneas circunscritas alrededor de un área o volumen dado y las obtenidas utilizando figuras rectilineas inscritas convergían ambas al mismo «último término». Señaló también que este proceso de paso al límite conduce a irracionales que no se pueden obtener como raíces de racionales. Pero estas ideas de Wallis y Gregory fueron ignoradas en su tiempo.

Los fundamentos del cálculo permancían oscuros. Se añadía a la

confusión el hecho de que los defensores del trabajo de Newton continuaban hablando de razones primeras y últimas, mientras que los seguidores de Leibniz utilizaban las cantidades no nulas infinitamente pequeñas. Muchos de los matemáticos ingleses, quizás porque en lo principal estaban todavía ligados al rigor de la geometría griega, recelaban de todo el trabajo sobre el cálculo. Así, el siglo terminaba con el cálculo en un estado de confusión.

Bibliografía

Armitage, A.: *Edmond Halley*, Thomas Nelson and Sons, 1966.

Auger, L.: *Un Savant Méconnu: Gilles Persone de Roberval (1602-1675)*, A. Blanchard, 1962.

Ball, W. W. R.: *A Short Account of the History of Mathematics*, Dover (reimpresión), 1960, pp. 309-370.

Baron, Margaret E.: *The Origins of Infinitesimal Calculus*, Pergamon Press, 1969.

Bell, Arthur E.; *Newtonian Science*, Edward Arnold, 1961.

Boyer, Carl B.: *The Concepts of the Calculus*, Dover (reimpresión), 1949.

— : *Historia de la Matemática*, Madrid, Alianza Editorial, 1986.

Brewster, David: *Memoirs of the Life, Writings and Discoveries of Sir Isaac Newton*, 2 vols., 1885, Johnson Reprint Corp., 1965.

Cajori, Florian: *A History of the Conceptions of Limits and Fluxions in Great Britain from Newton to Woodhouse*, Open Court, 1919.

— : *A History of Mathematics*, Macmillan, 1919, 2nd ed., pp. 181-220.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, 2.ª ed., B. G. Teubner, 1990 y 1898, vol. 2, pp. 821-922; vol. 3, pp. 150-316.

Child, J. M.: *The Geometrical Lectures of Isaac Barrow*, Open Court, 1916.

— : *The Early Mathematical Manuscript of Leibniz*, Open Court, 1920.

Cohen, I. B.: *Isaac Newton's Papers and Letters on Natural Philosophy*, Harvard University Press, 1958.

Coolidge, Julian L.: *The Mathematics of Great Amateurs*, Dover (reimpresión), 1963, caps. 7, 11 y 12.

De Morgan, Augustus: *Essays on the Life and Work of Newton*, Open Court, 1914.

Fermat, Pierre de: *OEuvres*, Gauthier-Villars, 1891-1912, vol. 1, pp. 139-179; vol. 3, pp. 121-156.

Gibson, G. A.: «James Gregory's Mathematical Work», *Proc. Edinburgh Math. Soc.*, 41, 1922-1923, 2-25.

Huygens, C.: *OEuvres complètes*, 22 vols., Societé Hollandaise des Sciences, Nyhoff, 1888-1950.

Leibniz, G. W.: *OEuvres*, Firmin-Didot, 1859-1875.
— : *Mathematische Schriften*, ed. C. I. Gerhardt, 7 vols., Ascher-Schmidt, 1849-1863. Reimpreso por Georg Olms, 1962.
More, Louis T.: *Isaac Newton*, Dover (reimpresión), 1962.
Montucla, J. F.: *Histoire des mathématiques*, Albert Blanchard (reimpresión), 1960, vol. 2, pp. 102-177, 348-403; vol. 3, pp. 102-138.
Newton, Sir Isaac: *The Mathematical Works*, ed. D. T. Whiteside, 2 vols., Johnson Reprint Corp., 1964-1967. El vol. 1 contiene la traducciones (al inglés) de los tres artículos básicos sobre el cálculo.
— : *Mathematical Papers*, ed. D. T. Whiteside, 4 vols., Cambridge University Press, 1967-1971.
— : *Mathematical Principles of Natural Philosophy*, ed. Florian Cajori, 3rd ed., University of California Press, 1946. *Principios matemáticos de la filosofía natural*, 2 vol., Madrid, Alianza Editorial, 1987.
— : *Opticks*, Dover (reimpresión), 1952. *Optica*, Madrid, Alfaguara, 1977.
Pascal, B.: *OEuvres*, Hachette, 1914-1921. *Obras. 1*, Madrid, Alfaguara, 1981.
Scott, Joseph F.: *The Mathematical Work of John Wallis*, Oxford University Press, 1938.
— : *A History of Mathematics*, Taylor and Francis, 1958, caps. 10-11.
Smith, D. E.: *A Source Book in Mathematics*, Dover (reimpresión), 1959, pp. 605-626.
Struik, D. J.: *A Source Book in Mathematics, 1200-1800*, Harvard University Press, 1969, pp. 188-316, 324-328.
Thayer, H. S.: *Newton's Philosophy of Nature*, Hafner, 1953.
Turnbull, H. W.: *The Mathematical Discoveries of Newton*, Blackie and Son, 1945.
— : *James Gregory Tercentenary Memorial Volume*, Royal Society of Edinburgh, 1939.
Turnbull, H. W. y Scott, J. F.: *The Correspondence of Isaac Newton*, 4 vols., Cambridge University Press, 1959-1967.
Walker, Evelyn: *A Study of the Traité des indivisibles of Gilles Persone de Roberval*, Columbia University Press, 1932.
Wallis, John: *Opera Mathematica*, 3 vols., 1693-1699, Georg Olms (reimpresión), 1968.
Whiteside, Derek T.: «Patterns of Mathematical Thought in the Seventeenth Century», *Archive for History of Exact Sciences*, 1, 1961, pp. 179-388.
Wolf, Abraham: *A History of Science, Technology and Philosophy in the 16th and 17th Centuries*, 2nd ed., George Allen and Unwin, 1950, caps. 7-14.

Capítulo 18
LAS MATEMATICAS A PARTIR DE 1700

Habiendo considerado esas pocas cosas, toda la cuestión se reduce a la geometría pura, la cual es el objetivo de la física y la mecánica.

G. W. LEIBNIZ

1. La transformación de las matemáticas

Al comienzo del siglo XVII, Galileo todavía había experimentado la necesidad de discutir con el pasado. Al final del siglo, las matemáticas habían experimentado cambios tan amplios y radicales que nadie hubiera dejado de percibir la llegada de una nueva era.

Los matemáticos europeos produjeron mucho más entre, aproximadamente, 1550 y 1700 de lo que los griegos habían producido en casi diez siglos. Esto se explica fácilmente por el hecho de que, mientras las matemáticas en Grecia se habían cultivado sólo por unos pocos, la difusión de la educación en Europa, aunque en absoluto universal, promovió el desarrollo de matemáticos en Inglaterra, Francia, Alemania, Holanda e Italia. La invención de la imprenta permitió un amplio acceso no sólo a los trabajos griegos, sino a los resultados de los propios europeos, los cuales, ahora más fácilmente disponibles, sirvieron para estimular nuevos esfuerzos.

Pero el genio del siglo no se pone en evidencia únicamente por la expansión de la actividad. La variedad de nuevos campos abiertos en este breve período es impresionante. El crecimiento del álgebra como ciencia (porque el uso de coeficientes literales permitió en

buena parte la realización de demostraciones), así como la vasta ampliación de sus métodos y teoría, los comienzos de la geometría proyectiva y de la teoría de la probabilidad, la geometría analítica, el concepto de función y, sobre todo, el cálculo, fueron las principales innovaciones, cada una de ellas destinada a empequeñecer la realización por excelencia de los griegos —la geometría euclídea.

Además del desarrollo cuantitativo y de los nuevos caminos de exploración se produjo un intercambio completo de los papeles del álgebra y de la geometría. Los griegos habían favorecido la geometría porque era el único camino en el que podían conseguir rigor; incluso en el siglo XVII, los matemáticos se sentían obligados a justificar los métodos algebraicos con demostraciones geométricas. Se podría decir que hasta 1600 el cuerpo de las matemáticas era geométrico, con algunos apéndices algebraicos y trigonométricos. Después del trabajo de Descartes, Fermat y Wallis el álgebra se convirtió no sólo en un método efectivo para sus propios fines, sino también en un enfoque superior para la solución de problemas geométricos. La mayor efectividad de los métodos analíticos en el cálculo decidió la competición, y el álgebra se convirtió en la sustancia dominante de las matemáticas.

Fueron Wallis y Newton quienes vieron claramente que el álgebra proporcionaba un método superior. A diferencia de Descartes, quien consideraba al álgebra como sólo técnica, Wallis y Newton se dieron cuenta de que se trataba de algo de vital importancia. El trabajo de Desargues, Pascal y La Hire fue menospreciado y olvidado, y los métodos geométricos de Cavalieri, Gregorio de St. Vicent, Huygens y Barrow fueron reemplazados. La geometría pura fue eclipsada durante casi cien años convirtiéndose, como mucho, en una interpretación del álgebra y en una guía del pensamiento algebraico mediante la geometría de coordenadas. Es cierto que la excesiva reverencia hacia el trabajo geométrico de Newton en los *Principia*, reforzada por la enemistad contra los matemáticos continentales engendrada por la disputa entre Newton y Leibniz, provocó que los matemáticos ingleses persistieran en el desarrollo geométrico del cálculo. Pero sus contribuciones fueron triviales comparadas con lo que los continentales fueron capaces de obtener utilizando el enfoque analítico. Lo que era evidente en 1700 fue expresado explícitamente nada menos que por una autoridad como Euler quien, en su *Introductio in Analysi Infinitorum* (1748), alaba al álgebra como muy superior a los métodos sintéticos de los griegos.

Los matemáticos abandonaron el método geométrico con gran desgana. Según Henry Pemberton (1694-1771), quien editó la tercera edición de los *Principia* de Newton, éste no sólo expresó constantemente una gran admiración por los geómetras de Grecia, sino que se censuró a sí mismo por no seguirlos más de cerca de lo que lo hizo. En una carta a David Gregory (1661-1708), un sobrino de James Gregory, Newton señalaba que «el álgebra es el análisis de los chapuceros en matemáticas». Pero su propia *Arithmetica Universalis* de 1707 hizo tanto como cualquier otro trabajo para establecer la supremacía del álgebra. En esta obra establece la aritmética y el álgebra como la ciencia básica, admitiendo sólo la geometría allí donde hiciera las demostraciones más fáciles. Leibniz, también, notó la dominación creciente del álgebra y se sintió obligado a decir, en un ensayo no publicado [1]: «a menudo, los geómetras pueden demostrar en pocas palabras lo que es muy largo en el cálculo... el enfoque del álgebra está garantizado, pero no es mejor.»

Otro cambio, más sutil, en la naturaleza de las matemáticas había sido aceptado inconscientemente por los maestros. Hasta 1550 los conceptos en matemáticas eran idealizaciones inmediatas o abstracciones de la experiencia. Por entonces hicieron su aparición los números irracionales y negativos, y fueron ganando aceptación gradualmente. Cuando, además, irrumpieron en las matemáticas los números complejos, una amplia álgebra que utilizaba coeficientes literales y las nociones de derivada e integral, el campo quedó dominado por conceptos procedentes de partes recónditas de la mente humana. La noción de un cambio relativo instantáneo, en particular, aunque teniendo, por supuesto, una base intuitiva en el fenómeno físico de la velocidad es, sin embargo, bastante más una construcción intelectual y es también, cualitativamente, una contribución completamente diferente del triángulo matemático. Además de estas ideas, entraban en liza las cantidades infinitamente grandes, que los griegos habían evitado cuidadosamente, y las cantidades infinitamente pequeñas, que los griegos habían soslayado habilidosamente.

En otras palabras, los matemáticos estaban aportando conceptos, más que abstrayendo ideas del mundo real. Sin embargo, estos conceptos eran útiles en las investigaciones físicas (con la excepción de los números complejos, que todavía tenían que probar su utilidad)

---

[1] Couturat, L.: *Opuscules et fragments inédits de Leibniz*, 1903; reimpreso por Georg Olms, 1961, p. 181.

porque tenían algunos lazos con la realidad física. Los europeos se sentían incómodos con los nuevos tipos de números y las nociones del cálculo, sin percibir realmente la causa de su preocupación. A medida que esos conceptos se iban revelando más útiles en las aplicaciones fueron siendo aceptados, primero a regañadientes y luego en forma pasiva. La familiaridad no produjo menosprecio sino aceptación e incluso naturalidad. Después de 1700, cada vez más nociones, cada vez más apartadas de la naturaleza y surgiendo auténticamente de la mente humana, entrarían a formar parte de las matemáticas con menos reservas. Para la génesis de sus ideas, las matemáticas se desplazaron de las facultades sensoriales a las intelectuales.

La incorporación del cálculo al cuerpo de las matemáticas introdujo otro cambio en el concepto mismo de las matemáticas, que subvirtió el ideal propugnado por los clásicos griegos. Hemos señalado ya que la emergencia del álgebra y del cálculo había planteado el problema de los fundamentos lógicos de estas partes de las matemáticas, y que este problema no había sido resuelto. A todo lo largo del siglo, algunos matemáticos se habían sentido incómodos por el abandono de la demostración en el sentido deductivo, pero sus protestas se ahogaron en el uso y satisfacción crecientes con respecto al álgebra y el cálculo; hacia finales de siglo, los matemáticos habían abandonado virtualmente los requerimientos de conceptos claramente definidos y demostraciones deductivas. La construcción axiomática rigurosa dio paso a la inducción a partir de ejemplos particulares, a ideas intuitivas, a vagas evidencias geométricas y a razonamientos físicos. Como la demostración deductiva había sido el rasgo distintivo de las matemáticas, los matemáticos estaban, pues, abandonando el sello de su campo de actividad.

Retrospectivamente es fácil ver por qué se vieron obligados a ello. Mientras los matemáticos obtuvieron sus conceptos de la experiencia inmediata, era posible definir estos conceptos y seleccionar los axiomas necesarios —aunque, además, las bases lógicas de la teoría de los enteros que presentó Euclides en los libros VII al IX de los *Elementos* eran deplorablemente deficientes—. Pero cuando introdujeron conceptos que ya no idealizaban experiencias inmediatas, como los números irracionales, negativos, y complejos, y la derivada y la integral, no valoraron adecuadamente el hecho de que estos conceptos eran de un carácter diferente y por ello no llegaron a darse cuenta de que se necesita una base para el desarrollo axiomático que fuera diferente de las verdades evidentes por sí mismas.

Es cierto, sin embargo, que los nuevos conceptos eran bastante más sutiles que los antiguos y que, como sabemos ahora, no podría haberse edificado fácilmente la base axiomática adecuada.

¿Cómo podían los matemáticos críticos, buenos conocedores de las matemáticas griegas, estar satisfechos actuando sobre una base heurística? Estaban preocupados por problemas importantes de la ciencia, acuciantes en algunos casos, y las matemáticas que utilizaban permitían manejar esos problemas. Más que buscar una comprensión total de las nuevas creaciones o intentar erigir la estructura deductiva requerida, justificaban su conciencia mediante sus éxitos. Un recurso ocasional a doctrinas filosóficas, o místicas, bastaba para encubrir algunas dificultades de manera que éstas dejaran de ser visibles.

Un objetivo nuevo, en particular, caracteriza las matemáticas del siglo XVII y siguientes —la generalidad de los métodos y de los resultados—. Ya hemos señalado el valor que se otorgaba a la generalidad del método: Vieta en su introducción de los coeficientes literales, los geómetras proyectivos, Fermat y Descartes en la exploración de curvas, y Newton y Leibniz en el tratamiento de las funciones. Por lo que se refiere a la generalidad de los resultados, los logros eran limitados. Muchos eran sólo afirmaciones, como la de que toda ecuación polinómica de grado n tiene n raíces, o que toda ecuación de segundo grado en x e y es una cónica. Los métodos matemáticos y la notación eran todavía demasiado limitados para permitir el establecimiento de resultados generales, pero esto se convirtió en un objetivo de los esfuerzos matemáticos.

2. Las matemáticas y la ciencia

Desde los tiempos de la Grecia clásica, las matemáticas se habían valorado sobre todo por su papel en la investigación de la naturaleza. La astronomía y la música estaban ligadas constantemente a las matemáticas, y la mecánica y la óptica eran, ciertamente, matemáticas. Sin embargo, la relación de las matemáticas con la ciencia se alteró de diversas formas debido al trabajo realizado en el siglo XVII. En primer lugar, porque la ciencia, que se estaba desarrollando enormemente, había sido dirigida por Galileo hacia la utilización de axiomas cuantitativos y deducciones matemáticas (cap. 16, sec. 3); la actividad matemática que estaba inspirada directamente por la ciencia se convirtió en dominante.

Además, la recomendación de Galileo de buscar la descripción matemática en lugar de la explicación causal condujo a la aceptación de conceptos tales como el de fuerza de gravitación. Esta fuerza y las leyes del movimiento fueron la base completa del sistema de Newton de la mecánica. Como el único conocimiento seguro sobre la gravitación era matemático, las matemáticas se convirtieron en la sustancia de las teorías científicas. El rebelde siglo XVII encontró un mundo cualitativo cuyo estudio se encontraba apoyado por las abstracciones matemáticas, y legó un mundo cuantitativo, matemático, que incluía bajo sus leyes matemáticas la concreción del mundo físico.

En tercer lugar, aunque los griegos había utilizado libremente las matemáticas en su ciencia, mientras bastaran las bases euclídeas para las matemáticas, existía una distinción profunda entre éstas y aquélla. Tanto Platón como Aristóteles distinguieron una de otra (cap. 3, sec. 10 y cap. 7, sec. 3), aunque de diferentes formas, y Arquímedes es especialmente claro acerca de lo que se establece matemáticamente y lo que se conoce físicamente. Sin embargo, a medida que se extendía el campo de las matemáticas, y que los matemáticos no sólo se basaban en significados físicos para entender sus conceptos, sino que aceptaban los razonamientos matemáticos porque proporcionaban profundos resultados físicos, la frontera entre las matemáticas y la ciencia se hizo borrosa. Paradójicamente, a medida que la ciencia iba basándose cada vez más en las matemáticas para producir sus conclusiones físicas, las matemáticas fueron basándose cada vez más en los resultados científicos para justificar sus propios procedimientos.

El resultado de esta interdependencia fue una virtual fusión de las matemáticas y de vastas áreas de la ciencia. Lo que constituía la brújula de las matemáticas, tal como se entendía en el siglo XVII, puede verse en el *Cursus seu Mundus Mathematicus* (El Curso o el Mundo de las Matemáticas) de Claude-François Milliet Deschales (1612-1678), publicado en 1674 y, en una edición ampliada, en 1690. Además de aritmética, trigonometría y logaritmos, trata de geometría práctica, mecánica, estática, geografía, magnetismo, ingeniería civil, carpintería, talla de piedras, construcción militar, hidrostática, movimiento de fluidos, hidráulica, construcción de barcos, óptica, perspectiva, música, diseño de armas de fuego y cañones, el astrolabio, relojes de sol, astronomía, el cálculo del calendario y el horóscopo. Finalmente, incluye álgebra, la teoría de los indivisibles, la teoría de las cónicas y curvas especiales como la cuadratiz y la espiral. Esta obra fue popular y estimada. Aunque en la inclusión de

algunos temas refleja los intereses del Renacimiento, presenta en su conjunto un cuadro razonable del mundo de las matemáticas del siglo XVII e incluso del XVIII.

Se podría esperar que los matemáticos hubieran estado preocupados por preservar la identidad de su dominio. Pero además del hecho de que estaban obligados a depender de significados y resultados físicos para sostener sus razonamientos, los más grandes entre los que contribuyeron con aportaciones a las matemáticas en el siglo XVII (y XVIII) fueron, o bien primariamente científicos o al menos preocupados por igual por los dos campos. Descartes, Huygens y Newton, por ejemplo, fueron más grandes físicos que matemáticos. Pascal, Fermat y Leibniz fueron activos en física. De hecho, sería difícil nombrar un matemático extraordinario de ese siglo que no hubiera estado vivamente interesado por la ciencia. Como consecuencia, estos hombres no quisieron, ni intentaron, hacer ninguna distinción entre los dos campos. Descartes dice, en sus *Reglas para la dirección del espíritu,* que las matemáticas son la ciencia del orden y la medida e incluye, además del álgebra y la geometría, la astronomía, la música, la óptica y la mecánica. Newton dice en sus *Principia:* «En matemáticas tenemos que investigar las cantidades de fuerzas, con su proporción consecuente, con cualesquiera condiciones supuestas; entonces, cuando entramos en la física, comparamos estas proporciones con los fenómenos de la naturaleza...» Aquí la física se refiere a la experimentación y a la observación. Las matemáticas de Newton podrían considerarse como la física matemática actual.

3. La comunicación entre los matemáticos

Hasta aproximadamente 1550, las matemáticas habían sido creadas por investigadores aislados o por pequeños grupos encabezados por uno o dos directores prominentes. Los resultados se comunicaban oralmente o se escribían ocasionalmente en textos —que, sin embargo, eran manuscritos—. Como las copias tenían que hacerse también a mano, éstas eran escasas. En el siglo XVII los libros impresos habían llegado a ser, de alguna manera, corrientes, aunque tampoco esta mejora difundió el conocimiento tan ampliamente como podría haberse pensado. Como el mercado para la matemática avanzada era pequeño, los editores tenían que establecer precios elevados. Los buenos impresores eran escasos. La publicación era segui-

da, habitualmente, por ataques a los autores por parte de oponentes no demasiado escrupulosos; no era difícil para tales críticos encontrar bases para el ataque, especialmente porque el álgebra y el cálculo no estaban en absoluto fundamentados lógicamente. Los libros, en cualquier caso, no eran habitualmente el vehículo de las nuevas creaciones, porque resultados significativos no garantizaban una publicación del tamaño de un libro.

Como consecuencia, muchos matemáticos se limitaron a escribir cartas a amigos en las que relataban sus descubrimientos. Temiendo que esas cartas llegaran a manos de quienes pudieran aprovecharse de tales documentos no oficiales, los escritores ponían sus resultados a menudo en forma cifrada o mediante anagramas, los cuales podían ser descifrados en caso necesario.

A medida que se incorporaban a la creación matemática más investigadores, el deseo de intercambio de información y el estímulo por reunirse con gente con los mismos intereses intelectuales se concretó en la formación de sociedades científicas o academias. En 1601 jóvenes nobles fundaron en Roma la Accademia dei Lincei (de los linces); duró treinta años. Galileo ingresó como miembro en 1611. Otra sociedad italiana, la Accademia del Cimento (Academia de Experimentos) fue fundada en Florencia en 1657 como una organización formal de investigadores que se habían estado reuniendo en un laboratorio fundado por dos miembros de la familia Medici aproximadamente diez años antes. Esta academia contó entre sus miembros con Vincenzo Viviani (1622-1703) y con Torricelli, ambos discípulos de Galileo. Desgraciadamente, la sociedad se deshizo en 1667. En Francia, Desargues, Descartes, Gassendi, Fermat y Pascal, entre otros, mantuvieron reuniones privadamente, bajo la dirección de Mersenne, desde 1630 en adelante. Este grupo fue distinguido por Luis XIV, fundando con él la Académie Royale des Sciences en 1666, y proporcionando apoyo a sus miembros. Paralelamente a lo que había sucedido en Francia, un grupo inglés centrado alrededor de John Wallis comenzó en 1645 a mantener reuniones en el Gresham College, en Londres. Estos investigadores se ocuparon sobre todo de matemáticas y astronomía. A este grupo Carlos II le concedió unos estatutos formales en 1662, adoptando el nombre de Royal Society of London for the Promotion of Natural Knowledge (Sociedad Real de Londres para la Promoción del Conocimiento Natural). Esta sociedad se dedicó a hacer útiles las matemáticas y la ciencia, considerando como temas de interés el teñido, el acuñado,

la fabricación de armas, el refinado de los metales y la estadística de poblaciones. Finalmente, la Academia de Ciencias de Berlín, por la que Leibniz había abogado durante algunos años, fue abierta en 1700, con Leibniz como primer presidente. En Rusia, Pedro el Grande fundó la Academia de Ciencias de San Petersburgo en 1724.

Las academias fueron importantes no sólo por hacer posible el contacto directo y el intercambio de ideas, sino porque apoyaron la existencia de revistas. La primera de las revistas científicas, aunque no fue patrocinada por una academia, fue el *Journal des Sçavans* o *Journal des Savants*, que comenzó su publicación en 1665. Esta revista y las *Philosophical Transactions of the Royal Society*, que comenzó a publicarse en el mismo año, fueron las primeras revistas que incluyeron artículos matemáticos y científicos. La Académie des Sciences francesa comenzó la publicación *Histoire de l'Académie Royale des Sciences avec les Mémoires de Mathématique et Physique* (Historia de la Real Academia de Ciencias con las Memorias de Matemáticas y Física). También publicó las *Mémoires de Mathématique et de Phisique Présentés a l'Académie Royale des Sciences par Divers Sçavans et Lus dans ses Assemblées* (Memorias de Matemáticas y Física presentadas a la Real Academia de Ciencias por Diversos Sabios y Leídas en sus Asambleas), también conocidas como las *Mémoires des Savants Etrangers*. Otra de las primeras revistas científicas, la *Acta Eruditorum*, comenzó en 1682 y, como estaba publicada en latín, pronto adquirió una difusión internacional. La Academia de Ciencias de Berlín patrocinó la *Histoire de l'Académie Royale des Sciences et Belles-lettres* (cuyo título durante muchos años fue el de *Miscellanea Berolinensia*).

Las academias y sus revistas abrieron nuevas puertas a la comunicación científica; éstas y revistas posteriores se convirtieron en el medio aceptado para la publicación de las nuevas investigaciones. Las academias propiciaron la investigación en cuanto que la mayor parte de ellas apoyaron económicamente a investigadores. Por ejemplo, Euler fue patrocinado por la Academia de Berlín desde 1741 hasta 1766 y Lagrange desde 1766 hasta 1787. La Academia de San Petersburgo apoyó a Daniel y a Nicholas Bernoulli en varias ocasiones, y a Euler desde 1727 hasta 1741, y otra vez desde 1766 hasta su muerte en 1783. La fundación de las academias por los gobiernos europeos marca también la entrada oficial de los gobiernos en el área de la ciencia y el apoyo a la misma. La utilidad de la ciencia había sido reconocida.

Las instituciones de las que cualquiera en la actualidad hubiera esperado que jugaran el papel más importante en la creación y difusión del conocimiento —las universidades— fueron ineficaces. Eran conservadoras y dogmáticas, controladas por las religiones oficiales de los países respectivos y muy lentas incorporando los conocimientos nuevos. En su conjunto, enseñaban sólo un poco de aritmética, álgebra y geometría. Aunque había algunos matemáticos en la universidad de Cambridge en el siglo XVI, desde 1600 hasta 1630 no hubo ninguno. De hecho, en Inglaterra a principios del siglo XVII las matemáticas no formaban parte del *curriculum*. Estaban consideradas como algo demoníaco. Wallis, que había nacido en 1616, dice de la educación corriente durante su infancia que «las matemáticas en aquel tiempo se consideraban raramente entre nosotros como algo académico; más bien se miraban como algo mecánico —un asunto de comerciantes—». Fue a la universidad de Cambridge y estudió matemáticas allí, pero aprendió bastante más estudiando independientemente. Aunque preparado como para ser profesor de matemáticas, se fue a Cambridge «porque ese estudio había muerto allí, y no se abría ningún porvenir para una profesor de esa materia».

Las primeras cátedras de matemáticas se fundaron en Oxford en 1619, y más tarde en Cambridge. Antes de eso, había habido sólo profesores de baja categoría. La cátedra Lucasiana en Cambridge, que Barrow fue el primero en ocupar, fue fundada en 1663. El mismo Wallis se hizo catedrático en Oxford en 1649 y mantuvo la cátedra hasta 1702. Un obstáculo para la contratación de profesores capaces fue que éstos tenían que tomar las órdenes sagradas, aunque se hicieron excepciones, como en el caso de Newton. Las universidades británicas (incluidas las de Londres, Glasgow y Edinburgo) tuvieron, en general, casi la misma historia: desde, aproximadamente, 1650 hasta 1750, fueron de alguna manera activas, pero declinaron en su actividad hasta, aproximadamente, 1825.

Las universidades francesas de los siglos XVII y XVIII fueron inactivas en matemáticas. Hasta finales del siglo XVIII, en que Napoleón fundó escuelas técnicas de primera categoría, no efectuaron ninguna contribución. También en las universidades alemanas la actividad matemática en esos dos siglos se mantuvo en un nivel bajo. Leibniz estuvo aislado y, como señalamos anteriormente, se lamentaba de las enseñanzas de las universidades. Los centros universitarios de Ginebra y Basilea, en Suiza, fueron las excepciones en el período que estamos considerando; se pudieron vanagloriar de los Bernoulli, Her-

mann, y otros. Las universidades italianas tuvieron alguna importancia en el siglo XVII pero perdieron terreno en el XVIII. Cuando se piensa que Pascal, Fermat, Descartes, Huygens y Leibniz no enseñaron nunca en ninguna universidad y que Kepler y Galileo, aunque enseñaron durante algún tiempo, fueron matemáticos de corte la mayor parte de su vida, se cae en la cuenta de lo relativamente poco importantes que fueron las universidades.

4. Las perspectivas para el siglo XVIII

Los avances enormes del siglo XVII en álgebra, geometría analítica y cálculo; el fuerte compromiso de las matemáticas con la ciencia, que proporcionó problemas profundos e interesantes; el entusiasmo producido por los sorprendentes éxitos de Newton en mecánica celeste y la mejora de las comunicaciones proporcionada por las academias y revistas apuntaba, todo ello, a desarrollos adicionales importantes y servía para crear unas expectativas inmensas sobre el futuro de las matemáticas.

Tenían que superarse algunos obstáculos. Las dudas sobre la solidez del cálculo, el distanciamiento entre los matemáticos ingleses y continentales, el bajo nivel de las instituciones educativas existentes y la incertidumbre sobre el apoyo a la profesión en matemáticas retenían a los jóvenes o futuros matemáticos. Sin embargo, el entusiasmo de los matemáticos era casi ilimitado. Tenían visiones de una tierra prometida y estaban ansiosos por presionar hacia adelante. Además, trabajaban en una atmósfera bastante más adecuada para la creación que en cualquier otra época, desde que la geometría de la Grecia clásica del 300 antes de Cristo imponía no sólo restricciones en el campo de las matemáticas, sino que imprimía un nivel de rigor a la matemática aceptable que impedía la creatividad. Los hombres del siglo XVII habían roto ambas ataduras. El progreso en matemáticas casi exige pasar por alto completamente los escrúpulos lógicos y, afortunadamente, los matemáticos se atrevieron ya a confiar en intuiciones y consideraciones físicas.

Bibliografía

Hahn, Roger: *The Anatomy of a Scientific Institution: The Paris Academy of Sciences, 1666-1803,* University of California Press, 1971.

Hall, A. Rupert: *The Scientific Revolution, 1500-1800,* Longmans, Green, 1954, cap. 7. *La Revolución científica, 1500-1700,* Barcelona, Crítica, 1985.

Hall, A. Rupert, y Marie Boas: *The Correspondence of Henry Oldenburg,* 4 vols., University of Wisconsin Press, 1968.

Hartley, Sir Harold: *The Royal Society: Its Origins and Founders,* The Royal Society, 1960.

Ornstein, M.: *The Role of Scientific Societies in the Seventeeth Century,* University of Chicago Press, 1938.

Purver, Margery: *The Royal Society, Concept and Creation,* Massachusetts Institute of Technology Press, 1967.

Wolf, Abraham: *A History of Science, Technology and Philosophy in the 16th and 17th Centuries,* 2.ª ed., George Allen and Unwin, 1950, cap. 4.

Morris Kline

El pensamiento matemático de la Antigüedad a nuestros días, II

Versión española de:
Carlos Fernández Pérez y Alejandro Garciadiego

Coordinación y revisión de Jesús Hernández

Alianza
Editorial

INDICE

Capítulo 19
EL CALCULO INFINITESIMAL EN EL SIGLO XVIII

> Y así pasa que los matemáticos de este tiempo actúan como hombres de ciencia, empleando mucho más esfuerzo en aplicar sus principios que en comprenderlos.
>
> BERKELEY

1. Introducción

El mayor logro del siglo XVII fue el cálculo infinitesimal. De ese manantial brotaron nuevas e importantes ramas de la matemática: ecuaciones diferenciales, series, geometría diferencial, cálculo de variaciones, funciones de variable compleja y muchas otras. El germen de alguna de estas materias estaba ya presente en los trabajos de Newton y Leibniz, y en el siglo XVIII estuvo dedicado en buena medida al desarrollo de estas ramas del análisis. Pero antes de que esto pudiese llevarse a cabo hubo que desarrollar el propio cálculo infinitesimal, pues, si bien Newton y Leibniz habían creado los métodos básicos, restaba mucho por hacer. Había que identificar como tales, o crearlas, muchas nuevas funciones de una variable así como funciones de dos o más variables; había que extender las técnicas de derivación y de integración a ciertas funciones conocidas y a otras todavía por conocer y quedaban por establecer los fundamentos lógicos del cálculo infinitesimal. El primer objetivo consistió en ampliar la materia objeto del cálculo infinitesimal y a ello están dedicados el presente capítulo y el próximo.

Los matemáticos del siglo XVIII extendieron el cálculo infinitesi-

mal y fundaron nuevas ramas del análisis, encontrándose en el proceso con los sufrimientos, los errores, las imperfecciones y la confusión de todo proceso creativo. Elaboraron un tratamiento puramente formal del cálculo infinitesimal y de las ramas del análisis de él resultantes. Su habilidad técnica fue insuperable, aunque no fue guiada por un elaborado pensamiento matemático sino por agudas percepciones de carácter físico e intuitivo. Estos esfuerzos formales resistieron la prueba de posteriores exámenes críticos y dieron lugar a grandes líneas de pensamiento. La conquista de nuevos dominios de la matemática tiene algo de las conquistas militares: ataques audaces en el territorio enemigo permiten capturar plazas fuertes y, después, estas incursiones han de ser seguidas y apoyadas por operaciones más amplias, profundas y cautelosas a fin de asegurar lo que sólo había sido alcanzado inseguramente a manera de ensayo.

Para apreciar el trabajo y los argumentos de los pensadores del siglo XVIII será útil tener presente que ellos no distinguían entre álgebra y análisis. Al no apreciar la necesidad del concepto de límite y, en consecuencia, los problemas que se introducían por el uso de series infinitas, contemplaban el cálculo infinitesimal, de un modo ingenuo, como una extensión del álgebra.

La figura clave en la matemática del siglo XVIII es Leonhard Euler (1707-83), físico teórico preeminente del siglo y hombre que hay que situar a la altura de Arquímedes, Newton y Gauss. Nacido cerca de Basilea de un padre pastor calvinsita, que quería que estudiase teología, ingresó en la universidad de esa ciudad, completando sus estudios a la edad de quince años. En Basilea estudió matemáticas con Jean Bernoulli; decidió dedicarse a esta ciencia y comenzó a publicar a la edad de dieciocho años, ganando a los diecinueve un premio de la Academia de Ciencias francesa por un trabajo sobre la arboladura de buques. Gracias a los hijos de Jean Bernoulli, Nicolaus (1695-1726) y Daniel (1700-82), consiguió un puesto en la Academia de San Petersburgo, en Rusia, comenzando como ayudante de Daniel Bernoulli y sucediéndole pronto como profesor. Aunque Euler pasó unos años difíciles (1733-41) bajo un gobierno autocrático, llevó a cabo una cantidad asombrosa de investigaciones cuyos resultados aparecieron en artículos publicados por la Academia de San Petersburgo. También colaboró con el gobierno ruso en numerosos problemas físicos. En 1741, invitado por Federico *el Grande*, se trasladó a Berlín, donde permaneció hasta 1766. A lo largo de este período, Euler impartió lecciones a la princesa de Anhalt-Des-

sau, sobrina del rey de Prusia; estas lecciones, sobre diversos temas —matemáticas, astronomía, física, filosofía y religión—, fueron publicadas más tarde como las *Cartas a una princesa alemana* y todavía se leen con placer. A petición de Federico *el Grande*, Euler trabajó sobre problemas de seguros así como diseño de canales y obras hidráulicas. Durante su estancia de veinticinco años en Berlín, también envió cientos de artículos a la Academia de San Petersburgo y la asesoró en sus asuntos.

En 1766, a petición de Catalina *la Grande*, regresó a Rusia, aunque temiendo los efectos del riguroso clima sobre su debilitada vista (había perdido la vista de un ojo en 1735); en efecto, se volvió ciego al poco de llegar a Rusia, permaneciendo los últimos diecisiete años de su vida totalmente privado de visión. No fueron por ello menos fructíferos esos años que los precedentes; Euler tenía una memoria prodigiosa; recordaba las fórmulas de trigonometría y de análisis así como las potencias, hasta la sexta, de los cien primeros números primos, por no hablar de innumerables poemas y de la *Eneida* entera. Su memoria era tan impresionante que podía realizar mentalmente cálculos que otros matemáticos competentes realizaban con dificultad sobre el papel.

La productividad matemática de Euler es increíble; sus principales campos de interés fueron el cálculo infinitesimal, las ecuaciones diferenciales, la geometría analítica y diferencial de curvas y superficies, la teoría de números, las series y el cálculo de variaciones, aplicando todo ello a todos los dominios de la física; él fue quien creó la mecánica analítica (en contraposición a la antigua mecánica geométrica) y la mecánica de los cuerpos rígidos; calculó el efecto de perturbación de los cuerpos celestes sobre la órbita de un planeta, así como las trayectorias de proyectiles en medios con rozamiento. Su teoría de las mareas y sus trabajos sobre diseño y velamen de buques contribuyeron a mejorar la navegación; en este dominio, su *Scientia Navalis* (1749) y la *Théorie complète de la construction et de la manœuvre des vaisseaux* (1773) son obras sobresalientes. Investigó el pandeo de vigas y calculó la carga de seguridad de una columna. En acústica, estudió la propagación del sonido y la consonancia y disonancia musical. Sus tres volúmenes sobre instrumentos ópticos contribuyeron al diseño de telescopios y microscopios; fue también el primero en tratar analíticamente las vibraciones de la luz y en deducir la ecuación del movimiento teniendo en cuenta la dependencia de la elasticidad y la densidad del éter, obteniendo mu-

chos resultados sobre refracción y dispersión de la luz. En el tema de la luz, fue el único físico del siglo XVIII que apoyó la teoría ondulatoria frente a la corpuscular. Las ecuaciones diferenciales fundamentales del movimiento de un fluido ideal le pertenecen, y las aplicó al flujo de la sangre en el cuerpo humano. En la teoría del calor, contempló éste, con Daniel Bernoulli, como una oscilación de moléculas, ganando un premio en 1783 con su *Ensayo sobre el fuego*. También le interesaron la química, la geografía y la cartografía, realizando un mapa de Rusia. Se decía que las aplicaciones eran una excusa para sus investigaciones matemáticas, pero no cabe duda que gustaba de ambas.

Euler escribió textos sobre mecánica, álgebra, análisis matemático, geometría diferencial y analítica y sobre cálculo de variaciones que fueron obras clásicas por más de cien años. En este capítulo nos ocuparemos de varias de ellas: los dos volúmenes de la *Introductio in Analysin Infinitorum* (1748), que constituye la primera exposición unificada del cálculo infinitesimal y el análisis elemental; la obra más amplia *Institutiones Calculi Differentialis* (1755), y los tres volúmenes de *Institutiones Calculi Integralis* (1768-70); todas ellas obras señeras. Los libros de Euler contenían todos algunas características muy originales; su mecánica, como se ha dicho, estaba basada en métodos analíticos más bien que geométricos; fue el primero en dar un tratamiento con entidad del cálculo de variaciones. Aparte de los textos, Euler publicó artículos de investigación originales de gran calidad a un ritmo de aproximadamente ochocientas páginas al año durante la mayor parte de los años de su vida; la calidad de estos artículos puede juzgarse por el hecho de que ganó tantos premios por ellos que las correspondientes dotaciones constituyeron un complemento casi regular de sus ingresos. Algunos de los libros, así como cuatrocientos de sus artículos de investigación, los escribió después de volverse completamente ciego. Cuando se complete la edición en curso de sus obras completas comprenderá setenta y cuatro volúmenes.

A diferencia de Descartes o Newton antes que él o de Cauchy después que él, Euler no inició nuevas ramas de la matemática, pero nadie fue más prolífico ni más diestro en utilizarla; nadie llegó a dominar y utilizar los recursos del álgebra, la geometría y el análisis para obtener tantos resultados admirables. Euler tuvo una magnífica inventiva metodológica y una gran habilidad técnica; nos topamos con su nombre en todas las ramas de la matemática: hay fórmulas

de Euler, polinomios de Euler, constantes de Euler, integrales eulerianas y líneas de Euler.

Podría pensarse que sólo pudo mantener tal volumen de actividad a costa de todos los demás intereses; pero Euler se casó y tuvo trece hijos, estando siempre atento al bienestar de su familia; educó a sus hijos y nietos, construyendo juegos científicos para ellos y pasando tardes leyéndoles la Biblia. También era aficionado a opinar sobre cuestiones filosóficas, aunque aquí descubrió su único punto débil y recibió por ello frecuentes pullas de Voltaire; en una ocasión se vio forzado a reconocer que nunca había estudiado filosofía y lamentó haber creído que se podía comprender dicha materia sin haberla estudiado; pero el ánimo de Euler para las disputas filosóficas no disminuyó y continuó empeñandose en ellas; incluso se divertía con las mordaces críticas que recibía de Voltaire.

Rodeado de un respeto universal —bien merecido por la nobleza de su carácter— pudo, al final de su vida, considerar como discípulos suyos a todos los matemáticos de Europa. El 7 de septiembre de 1783, después de charlar sobre los asuntos del día, los Montgolfiers [1] y el descubrimiento de Urano, «cesó de calcular y de vivir», según las muy citadas palabras de J. A. N. C. de Condorcet.

2. El concepto de función

Como hemos visto, durante el siglo XVII se introdujeron y utilizaron tanto el concepto de función como las funciones algebraicas y trascendentes más simples. A medida que Leibniz, Jacques y Jean Bernoulli, L'Hospital, Huygens y Pierre Varignon (1654-1722) abordaban problemas como el movimiento del péndulo, el perfil de una cuerda suspendida de dos puntos fijos, el movimiento a lo largo de trayectorias curvilíneas, el movimiento con rumbo constante sobre una esfera (la loxódroma), evolutas e involutas de curvas, cáusticas que aparecen en reflexión y refracción de la luz y la trayectoria de un punto de una curva que gira sobre otra, no sólo empleaban las funciones ya conocidas sino que llegaban a formas más complicadas de funciones elementales. Como consecuencia de estas investigaciones y del avance, en general, del cálculo infinitesimal, las funciones

[1] Los Montgolfiers eran dos hermanos que en 1783 lograron por vez primera ascender en un globo inflado con aire caliente.

elementales alcanzaron un nivel de conocimiento y desarrollo prácticamente equivalente al de hoy en día. Por ejemplo, la función logarítmica, originada como relación entre los términos de una progresión geométrica y una aritmética y que fue tratada en el siglo XVII como la serie resultante de la integración de $1/(1 + x)$ (cap. 17, sec. 2), fue introducida sobre una nueva base. El estudio de la función exponencial por Wallis, Newton, Leibniz y Jean Bernoulli mostró que la función logarítmica era la inversa de la exponencial, cuyas propiedades son relativamente simples. William Jones (1675-1749) dio, en 1742, una introducción sistemática a la función logarítmica de esa manera (cap. 13, sec. 2). Euler, en su *Introductio,* define ambas funciones como

$$e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right), \qquad \log x = n(x^{1/n} - 1).$$

También se sistematizó el estudio matemático de las funciones trigonométricas, de las que Newton y Leibniz dieron desarrollos en serie. Las fórmulas para las funciones de la suma y la diferencia de dos ángulos, tales como sen $(x + y)$ o sen $(x - y)$, se deben a muchas personas, entre las cuales están Jean Bernoulli y Thomas Fancet de Lagny (1660-1734); este último escribió un artículo sobre este tema en las *Mémoires* de la Academia de París en 1703. Frédéric-Christian Mayer (de quien se desconocen las fechas de nacimiento y muerte), uno de los primeros miembros de la Academia de Ciencias de San Petersburgo, derivó a continuación las fórmulas usuales de la trigonometría a partir de las fórmulas de la suma y la diferencia [2]. Finalmente, Euler, en un artículo premiado de 1748 sobre las anomalías en los movimientos de Júpiter y Saturno, dio el tratamiento sistemático completo de las funciones trigonométricas [3]. La periodicidad de estas funciones está clara en la *Introductio* (1748), en donde Euler introdujo también la medida en radianes de los ángulos [4].

El estudio de las funciones hiperbólicas comenzó cuando se observó que el área bajo una circunferencia estaba dada por $\int \sqrt{a^2 - x^2}\, dx$, mientras que el área bajo la hipérbola estaba dada por $\int \sqrt{x^2 - a^2}\, dx$. Como ambas difieren en un signo, y el área bajo la circunfenrencia

[2] *Com. Acad. Sci. Petrop.*, 2, 1727.
[3] *Opera, (2), 25, 45-157.*
[4] *Opera,* (1), 9, 217-329, 305-307.

se puede expresar mediante funciones trigonométricas (poniendo $x = a$ sen θ), mientras que el área bajo la hipérbola está relacionada con la función logarítmica, debería existir una relación que incluye números imaginarios entre las funciones trigonométricas y la función logarítmica; esta idea fue desarrollada por gran número de personas (ver sec. 3) hasta que, finalmente, J. H. Lambert estudió por extenso las funciones hiperbólicas [5].

El concepto de función había sido formulado por Jean Bernoulli. Euler, en el mismo comienzo de la *Introductio*, define una función como cualquier expresión analítica formada, de modo arbitrario, a partir de una cantidad variable y de constantes; incluye los polinomios, las series de potencias y las expresiones trigonométricas y logarítmicas; también define las funciones de varias variables. Euler comienza con la noción de función algebraica, en la que las operaciones que se hacen sobre la variable independiente son únicamente algebraicas, las cuales a su vez se dividen en dos clases: racionales, en las que intervienen solamente las cuatro operaciones elementales, y las irracionales, en las que intervienen raíces. A continuación, introduce las funciones trascendentes, a saber, las trigonométricas, la logarítmica, la exponencial, las potencias de exponente irracional y ciertas integrales.

La principal diferencia entre las funciones, escribe Euler, consiste en la combinación de variables y constantes que las componen. Así, añade, las funciones trascendentes se distinguen de las algebraicas en que aquéllas repiten un número infinito de veces las operaciones de estas últimas. Es decir las funciones trascendentes estarían dadas por series infinitas. Euler y sus contemporáneos no se planteaban la necesidad de considerar la validez de las expresiones obtenidas al aplicar infinitas veces las cuatro operaciones racionales.

Euler distinguía entre funciones implícitas y explícitas y entre funciones univalentes y multivalentes, siendo estas últimas las raíces de ecuaciones en dos variables de grado superior cuyos coeficientes son funciones de una variable. En este punto, señala, si una función tal como $\sqrt[3]{P}$, donde P es una función univalente, toma valores reales para valores reales del argumento, entonces se podrá incluir en la mayoría de las ocasiones entre las funciones univalentes. A partir de estas definiciones (que no están libres de contradicciones), Euler con-

[5] *Hist. de l'Acad. de Berlin*, 24, 1768, 327-354, pub. 1770 = *Opera Math.*, 2, 245-269.

sidera las funciones racionales enteras o polinomios; estas funciones con coeficientes reales se pueden descomponer, afirma Euler, en factores de primero y segundo grado con coeficientes reales (ver sec. 4 y cap. 25, sec. 2).

Por función continua, Euler, como Leibniz y otros pensadores del siglo XVIII, entendía una función especificada por una fórmula analítica; su término «continua» significa en realidad «analítica» para nosotros, excepto en lo que se refiere a una discontinuidad excepcional como en $y = 1/x$ [6]. También fueron identificadas otras funciones y las curvas que las representaban se calificaban de «mecánicas» o «de trazo libre».

La *Introductio* de Euler fue la primera obra en que se estableció el concepto de función como una noción básica sobre la que ordenar el material de los dos volúmenes de aquélla. Algo del espíritu de este libro puede extraerse de las observaciones de Euler sobre el desarrollo de funciones en series de potencias [7]. Afirma allí que cualquier función puede desarrollarse de ese modo, pero en seguida dice que «si alguien duda de que cualquier función puede desarrollarse así, la duda quedará desechada desarrollando de hecho la función. Sin embargo, con el fin de que la presente investigación abarque el dominio más amplio posible, además de las potencias enteras positivas de z, también se admitirán términos con exponentes arbitrarios. De este modo, es ciertamente evidente que cualquier función puede expresarse en la forma $Az^{\alpha} + Az^{\beta} + Cz^{\gamma} + Dz^{\delta} + \cdots$, donde los exponentes $\alpha, \beta, \gamma, \delta \cdots$ pueden ser números cualesquiera». Para Euler, la posibilidad de desarrollar en serie todas las funciones estaba confirmada por su propia experiencia y la de todos sus contemporáneos; y, de hecho, era cierto en aquellos tiempos que todas las funciones dadas por expresiones analíticas admitían un desarrollo en serie.

Aunque surgió una controversia acerca de la noción de función en relación con el problema de la cuerda vibrante (ver cap. 22) que llevó a Euler a generalizar su propia noción de lo que era una función, el concepto que predominó en el siglo XVIII fue todavía el de función dada por una única expresión analítica, finita o infinita. Así, Lagrange, en su *Théorie des fonctions analytiques* (1797), definía una

[6] En el volumen 2, capítulo 1 de su *Introductio*, Euler introduce funciones «discontinuas» o mixtas que requieren expresiones analíticas diferentes en diferentes dominios de la variable independiente, pero este concepto no desempeña ningún papel en su obra.

[7] *Opera*, (1), 8, cap. p. 74.

función de una o varias variables como cualquier expresión útil para el cálculo en que dichas variables intervenían de cualquier manera. En las *Leçons sur le calcul des fonctions* (1806), dice que las funciones representan distintas operaciones que han de realizarse sobre cantidades conocidas para obtener los valores de cantidades desconocidas, y que éstas son estrictamente sólo el último resultado del cálculo. En otras palabras, una función es una combinación de operaciones.

3. Técnicas de integración y cantidades complejas

El método básico para integrar funciones algebraicas, por mínimamente complicadas que fueran, y funciones trascendentes consistía en representar las funciones en serie e integrar término a término, técnica que fue introducida por Newton. Poco a poco, los matemáticos fueron desarrollando técnicas que permitían pasar de una forma cerrada a otra.

El uso del concepto de integral en el siglo XVIII fue bastante restringido. Newton había utilizado la derivada y la antiderivada, o integral indefinida, mientras que Leibniz puso el énfasis en las diferenciales y su suma. Jean Bernoulli, presumiblemente siguiendo a Leibniz, trató la integral como inversa de la diferencial, de modo que si $dy = f'(x)\,dx$, entonces $y = f(x)$. Es decir, la antiderivada de Newton se tomaba como integral, pero se utilizaba la diferencial en lugar de la derivada de Newton. De acuerdo con Bernoulli, el objeto del cálculo integral era encontrar, a partir de una relación dada entre *diferenciales* de variables, la relación existente entre las variables. Euler subrayó que la derivada era la razón entre las diferenciales evanescentes y dijo que el cálculo integral se ocupaba de hallar la propia función; sólo utilizó la idea de suma para la evaluación aproximada de integrales. En realidad, todos los matemáticos del siglo XVIII trataron la integral como inversa de la derivada o de la diferencial dy. La existencia de una integral nunca fue puesta en cuestión; se determinaba, por supuesto, explícitamente en la mayoría de aplicaciones que se realizaban en ese siglo, con lo que la cuestión no se planteaba.

Merece la pena considerar algunos ejemplos del desarrollo de las técnicas de integración. Para calcular

$$\int \frac{a^2\,\mathrm{dx}}{a^2 - x^2}$$

Jacques Bernoulli [8] había utilizado el cambio de variable

$$x = a\,\frac{b^2 - t^2}{b^2 + t^2},$$

que convierte la integral en

$$\int \frac{dt}{2at},$$

la cual se integra inmediatamente para dar una función logarítmica. Jean Bernoulli observó en 1702, y lo publicó en las *Mémoires* de la Academia de Ciencias de ese año [9], que

$$\frac{a^2}{a^2 - x^2} = \frac{a}{2}\left(\frac{1}{a+x} + \frac{1}{a-x}\right),$$

lo que permite una integración inmediata. Así surgió el método de descomposición en fracciones simples, método que también fue indicado independientemente por Leibniz en el *Acta Eruditorum* de 1702 [10].

Jean Bernoulli y Leibniz, en la correspondencia entre ambos, aplicaron dicho método a la integral

$$\int \frac{dx}{ax^2 + bx + c}.$$

Sin embargo, como los factores lineales de $ax^2 + bx + c$ pueden ser complejos, el método de descomposición en fracciones simples lleva a integrales de la forma

$$\int \frac{dx}{cx + d},$$

en las que d, al menos, es un número complejo. No obstante, tanto Leibniz como Jean Bernoulli realizaban la integración utilizando la regla del logaritmo, haciendo intervenir así los logaritmos de núme-

[8] *Acta Erud.*, 1969 = *Opera*, 2, 868-870.
[9] *Opera*, 1, 393-400.
[10] *Math. Schriften*, 5, 350-366.

ros complejos. A pesar de la confusión existente acerca de los números complejos, ninguno de ellos dudó en integrar de esa manera; Leibniz decía que la presencia de números complejos no importaba y Jean Bernoulli los empleó repetidamente. En un artículo publicado en 1702 [11], éste señalaba que, del mismo modo que $adz/(b^2 - z^2)$ se transforma por medio de la sustitución $z = b(t-1)/(t+1)$ en $adt/2bt$, la diferencial

$$\frac{dz}{b^2 + z^2}$$

se transforma por la sustitución $z = \sqrt{-1}\, b(t-1)/(t+1)$ en

$$\frac{-dt}{\sqrt{-1}\, 2bt},$$

y que esta última es la diferencial del logaritmo de un número complejo. Como la integral original conduce también a la función arco tangente, Bernoulli establecía así una relación entre las funciones trigonométricas y la logarítmica.

Sin embargo, estos resultados en seguida suscitaron una viva polémica acerca de la naturaleza de los logaritmos de números negativos y de números complejos. En su artículo de 1712 [12] y en un intercambio de cartas con Jean Bernoulli durante los años 1712-13, Leibniz afirmaba que los logaritmos de números negativos eran inexistentes (él decía imaginarios), mientras que Bernoulli intentaba probar que tenían que ser reales. El argumento de Leibniz era que los logaritmos positivos se utilizaban para números mayores que 1 y los logaritmos negativos para números entre 0 y 1, con lo que no podían existir logaritmos para los números negativos; además, si -1 tuviese logaritmo, el logaritmo de $\sqrt{-1}$ tendría que valer la mitad, pero era evidente que $\sqrt{-1}$ no podía tener logaritmo. Que Leibniz argumentase de este modo después de haber introducido los logaritmos de números complejos en integración resulta inexplicable. Por su parte, Bernoulli argüía que dado que

$$\frac{d(-x)}{-x} = \frac{dx}{x}, \tag{1}$$

[11] *Mém. de l'Acad. des Sci., Paris,* 1702, 289 ff. = *Opera,* 1, 393-400.

[12] *Acta Erud.*, 1712, 167-169 = *Math. Schriften,* 5, 387-389.

entonces $\log(-x) = \log x$, y como $\log 1 = 0$, lo mismo ocurre con $\log(-1)$; a ello replicó Leibniz señalando que $d(\log x) = dx/x$ sólo es válido para x positivos. Un segundo asalto de correspondencia y desacuerdo tuvo lugar entre Euler y Jean Bernoulli durante los años 1727-31. Bernoulli mantuvo su postura mientras que Euler mostró su desacuerdo con ella, aunque, a la vez, sin ofrecer una postura propia consistente.

La clarificación final de lo que es el logaritmo de un número complejo se hizo posible gracias a otros desarrollos análogos que tienen importancia por sí mismos y que llevan a la relación existente entre la función exponencial y las trigonométricas. En 1714 Roger Cotes (1682-1716) publicó [13] un teorema sobre números complejos que, en notación moderna, establece que

$$\sqrt{-1}\,\phi = \log_e(\cos\phi + \sqrt{-1}\,\text{sen}\,\phi). \qquad (2)$$

En una carta a Jean Bernoulli del 18 de octubre de 1740, Euler afirmaba que $y = 2\cos x$ e $y = e^{\sqrt{-1}x} + e^{-\sqrt{-1}x}$ eran ambas soluciones de la misma ecuación diferencial (que él identificó gracias a soluciones en serie) con lo que habían de ser iguales. Publicó esta observación en 1743 [14], a saber,

$$\cos s = \frac{e^{\sqrt{-1}s} + e^{-\sqrt{-1}s}}{2}, \qquad \text{sen}\, s = \frac{e^{\sqrt{-1}s} - e^{-\sqrt{-1}s}}{2\sqrt{-1}}. \qquad (3)$$

En 1748, Euler redescubrió el resultado (2) de Cotes, que también podría deducirse de (3).

Mientras se producían estos progresos, Abraham de Moivre (1667-1754), que abandonó Francia y se estableció en Londres cuando fue revocado el Edicto de Nantes que protegía a los hugonotes, obtuvo, al menos implícitamente, la fórmula que hoy lleva su nombre. En una nota de 1722, que utiliza un resultado ya publicado en 1707 [15], afirma que se puede obtener una relación entre x y t, que representan los senoversos de dos arcos (senver $\alpha = 1 - \cos\alpha$) que están en una razón de 1 a n, eliminando z de las dos ecuaciones

$$1 - 2z^n + z^{2n} = -2z^n t \quad \text{y} \quad 1 - 2z + z^2 = -2zx.$$

[13] *Phil. Trans.*, 29, 1714, 5-45.
[14] *Miscellanea Berolinensia*, 7, 1743, 172-192 = *Opera*, (1), 14, 138-155.
[15] *Phil. Trans.*, 25, 1707, 2368-2371.

En este resultado está implícita la fórmula de De Moivre, ya que si se pone $x = 1 - \cos\phi$, $t = 1 - \cos n\phi$, se obtiene

$$(\cos\phi \pm \sqrt{-1}\ \mathrm{sen}\ \phi)^n = \cos n\phi \pm \sqrt{-1}\ \mathrm{sen}\ n\phi. \qquad (4)$$

Para de Moivre, n era una entero positivo; en realidad, él nunca escribió este último resultado explícitamente; fue Euler quien dio la formulación final [16] y quien lo generalizó para todo número real n.

Para 1747, Euler disponía ya de la suficiente experiencia con las relaciones entre exponenciales, logaritmos y funciones trigonométricas como para obtener los resultados correctos sobre logaritmos de números complejos. En un artículo de 1749, titulado «De la controversia entre Messrs. Leibnitz et Bernoulli sur les logarithmes négatifs et imaginaires» [17], Euler se muestra en desacuerdo con el contraargumento de Leibniz de que $d(\log x) = dx/x$ sólo para x positivo. Según él, si la objeción de Leibniz fuese correcta quebraría los fundamentos de todo el análisis, a saber, que las reglas y operaciones son válidas sea cual sea la naturaleza de los objetos a los que se aplican aquéllas. Euler afirma que $d(\log x) = dx/x$ es correcta para valores positivos y negativos de x, pero añade que Bernoulli olvida que de la fórmula (1) de más arriba sólo se puede concluir que $\log(-x)$ y $\log(x)$ difieren en una constante. Esta constante ha de ser $\log(-1)$, ya que $\log(-x) = \log(-1 \cdot x) = \log(-1) + \log x$. En efecto, Bernoulli ha supuesto que $\log(-1) = 0$, pero hay que demostrarlo. Bernoulli había proporcionado otros argumentos a los que también respondió Euler. Por ejemplo, Bernoulli afirmaba que dado que $(-a)^2 = a^2$, entonces $\log(-a)^2 = \log a^2$, de donde $2\log(-a) = 2\log a$ y, por tanto, $\log(-a) = \log a$. Euler replica que dado que $(a\sqrt{-1})^4 = a^4$, se tiene que $\log a = \log(a\sqrt{-1}) = \log a + \log\sqrt{-1}$, con lo que, en este caso, cabe presumir que $\log\sqrt{-1}$ debería ser 0. Pero el propio Bernoulli, dice Euler, estableció en otro contexto que $\log\sqrt{-1} = \sqrt{-1}\,\pi/2$.

Leibniz había argumentado que dado que

$$\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots, \qquad (5)$$

[16] *Introductio*, cap. 8.

[17] *Hist. de l'Acad. de Berlin*, 5, 1749, 139-179, pub. 1751 = *Opera*, (1), 195-232.

entonces para $x = -2$

$$\log(-1) = -2 - \frac{4}{2} - \frac{8}{3} \cdots,$$

de donde se deduce al menos que log (− 1) no es 0 (de hecho, Leibniz había dicho que log (− 1) no existía). La respuesta de Euler a este argumento fue que de

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 \cdots,$$

se tiene para $x = -3$

$$-\frac{1}{2} = 1 + 3 + 9 + 27 + \cdots,$$

mientras que para $x = 1$ resulta

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots,$$

con lo que sumando miembro a miembro se obtiene

$$0 = 2 + 2 + 10 + 26 + \cdots.$$

En consecuencia, afirma Euler, el argumento basado en las series no prueba nada.

Después de refutar a Leibniz y Bernoulli, Euler da lo que, según los criterios actuales, es un argumento incorrecto. Escribe

$$x = e^y = \left(1 + \frac{y}{i}\right)^i,$$

en donde i es un número infinitamente grande [18]. Entonces

$$x^{1/i} = 1 + \frac{y}{i}$$

[18] Antes de 1777, Euler utilizó la letra i (por *infinitus*) para una cantidad infinitamente grande; después de esa fecha utlizó la i para $\sqrt{-1}$.

y, en consecuencia,

$$y = i\ (x^{1/i} - 1).$$

Como $x^{1/i}$ —«la raíz con exponente i infinitamente grande»— toma infinitos valores complejos, eso ocurre con y, y como $y = \log x$, lo mismo se puede afirmar de $\log x$. De hecho, Euler escribe [19]

$$x = a + b\ \sqrt{-1} = c\ (\cos\ \phi + \sqrt{-1}\ \text{sen}\ \phi).$$

Poniendo $c = e^C$ se obtiene

$$x = e^C\ (\cos\ \phi + i\ \text{sen}\ \phi) = e^C e^{\sqrt{-1}\ (\phi \pm 2\lambda\pi)}$$

y así

$$y = \log\ x = C + \ (\phi \pm 2\lambda\pi)\ \sqrt{-1}, \qquad (6)$$

donde λ es un entero positivo o cero. En consecuencia, afirma Euler, para los números reales positivos, sólo un valor del logaritmo es real siendo imaginarios todos los demás valores; para los números reales negativos y para los imaginarios, sin embargo, todos los valores del logaritmo son imaginarios. A pesar de esta brillante resolución del problema, el trabajo de Euler no fue aceptado. D'Alembert formuló argumentos de carácter metafísico, analítico y geométrico para mostrar que $\log\ (-1) = 0$.

4. Integrales elípticas

Después de haber logrado integrar algunas funciones racionales por el método de fracciones simples, Jean Bernoulli afirmó en las *Acta Eruditorum* de 1702 que la integral de cualquier función racional no implicaba más funciones trascendentes que las trigonométricas y la logarítmica. Como el denominador de una función racional es un polinomio en x de grado n, la validez de esa afirmación de-

---

[19] $a + b\sqrt{-1} = \sqrt{a^2 + b^2}\left(\dfrac{a}{\sqrt{a^2 + b^2}} + \sqrt{-1}\ \dfrac{b}{\sqrt{a^2 + b^2}}\right) = c\,(\cos\ \phi + i\ \text{sen}\ \phi).$

pendía de si cualquier polinomio con coeficientes reales podía expresarse como un producto de factores de primer y segundo grado con coeficientes reales. En su artículo de las *Acta* de 1702, Leibniz opinaba que ello no era posible y daba el ejemplo de $x^4 + a^4$. Señalaba que

$$x^4 + a^4 = (x^2 - a^2\sqrt{-1})\,(x^2 + a^2\sqrt{-1})$$
$$= (x + a\sqrt{\sqrt{-1}})\,(x - a\sqrt{\sqrt{-1}})$$
$$\times\,(x + a\sqrt{-\sqrt{-1}})\,(x - a\sqrt{-\sqrt{-1}})$$

y, según él, para ningún par de esos cuatro factores se verificaba que su producto fuese un factor cuadrático con coeficientes reales. Si hubiese sido capaz de expresar la raíz cuadrada de $\sqrt{-1}$ y de $-\sqrt{-1}$ como números complejos ordinarios, se hubiese apercibido de su error. Nicolaus Bernoulli (1687-1759), un sobrino de Jacques y Jean, indicó en las *Acta Eruditorum* de 1719 que $x^4 + a^4 = (a^2 + x^2)^2 - 2a^2x^2 = (a^2 + x^2 + ax\sqrt{2})\,(a^2 + x^2 - ax\sqrt{2})$, de donde se sigue que la función $1/(x^4 + a^4)$ se puede integrar en términos de funciones trigonométricas y de la logarítmica.

También se estudió la integración de funciones irracionales. Jacques Bernoulli y Leibniz se escribieron sobre este tema debido a que tales integrandos aparecían con frecuencia. En 1694 [20], Jacques estaba interesado en la elástica, el perfil que adopta una barra delgada cuando se ejercen fuerzas sobre ella —por ejemplo, en sus extremos—. Para cierto conjunto de condiciones en los extremos, encontró que la ecuación de la curva está dada por

$$dy = \frac{(x^2 + ab)\,dx}{\sqrt{a^4 - (x^2 + ab)^2}}\,;$$

expresión que no pudo integrar en términos de funciones elementales. En relación con este trabajo introdujo la lemniscata, cuya ecuación en coordenadas rectangulares es $(x^2 + y^2)^2 = a^2(x^2 - y^2)$ mientras que en coordenadas polares es $r^2 = a^2 \cos 2\theta$. Jacques Bernoulli intentó hallar la longitud de arco, que, desde el vértice a un punto arbitrario de la curva, está dada por

$$s = \int_0^r \frac{a^2}{\sqrt{a^2 - r^4}}\,dr,$$

[20] *Acta Erud.*, 1964, 262-276 = *Opera*, 2, 575-600.

y conjeturó que tampoco esta integral podía integrarse en términos de funciones elementales. Las tentativas del siglo XVII para rectificar la elipse, cuya longitud de arco tiene importancia en Astronomía, condujeron al problema de evaluar

$$s = a \int_0^t \frac{(1-k^2 t^2)\, dt}{\sqrt{(1-t^2)(1-k^2 t^2)}},$$

cuando la ecuación de la elipse está dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

y en el integrando $k = (a^2 - b^2)/a^2$ y $t = x/a$. El problema de hallar el período del péndulo simple condujo a la integral

$$T = 4\sqrt{l/g} \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-k^2 \operatorname{sen} \phi^2}}.$$

También aparecieron integrandos irracionales al calcular la longitud de arco de una hipérbola, de las funciones trigonométricas y de otras curvas. Estas integrales ya eran conocidas antes de 1700 y a lo largo del siglo XVIII continuaron apareciendo otras más con integrandos de esa clase. Así, Euler, en un tratamiento definitivo de la elástica realizado en el apéndice de libro de 1744 sobre cálculo de variaciones, obtuvo

$$dy = \frac{(\alpha + \beta x + \gamma x^2)\, dx}{\sqrt{a^4 - (\alpha + \beta x + \gamma x^2)^2}},$$

donde no importa el valor de las constantes. Como sus predecesores, Euler recurrió a desarrollos en serie a fin de obtener resultados físicos.

La clase de integrales que comprende los ejemplos anteriores se conoce como la de las integrales elípticas, proviniendo el nombre del cálculo de la longitud de arco de una elipse. En el siglo XVIII no se sabía, pero estas integrales no se pueden evaluar en términos de las funciones algebraicas, las circulares, la logarítmica o la exponencial [21].

Las primeras investigaciones sobre integrales elípticas estaban di-

---

[21] Esto fue demostrado por Liouville (*Jour. de l'Ecole Poly.*, 14, 1833, 124-193).

rigidas no tanto a evaluarlas como a intentar reducir las más complicadas a las que surgen al rectificar la elipse y la hipérbola. La razón de este enfoque estriba en que desde el punto de vista geométrico, que era el que primaba en la época, las integrales para los arcos elípticos e hiperbólicos parecían ser las más simples. Se inició un nuevo punto de vista con la observación de que la ecuación diferencial

$$f(x)\,dx = \pm f(y)\,dy, \tag{7}$$

donde $\int f(x)\,dx$ es una función logarítmica o una función trigonométrica inversa, posee una integral que es una función algebraica de x e y; es decir, a pesar de que es imposible encontrar una integral algebraica de la propia $f(x)\,dx$, sí se puede encontrar una integral algebraica de la suma o la diferencia de dos de esas diferenciales. Jean Bernoulli se preguntó entonces si esa propiedad podría ser cierta para integrales de otras funciones distintas de los logaritmos y las funciones trigonométricas inversas [22]. El mismo había descubierto en 1698 que la diferencia de dos arcos de la parábola cúbica ($y = x^3$) es integrable, resultado que obtuvo accidentalmente y que consideró de lo más elegante. Planteó entonces el problema más general de encontrar arcos de parábolas, elipses e hipérbolas (de orden superior) cuya suma o diferencia fuese igual a una cantidad rectilínea, y afirmó que ello ocurría para curvas parabólicas de la forma $a^m y^p = b^n x^q$, $m + p = n + q$, aunque no dio ninguna demostración.

El conde Giulio Carlo de' Toschi di Fagnano (1682-1766), un matemático aficionado, comenzó en 1714 a ocuparse de estos problemas [23]. Consideró las curvas $y = (2/m + 2)\, x^{(m+2)}/\, a^{m/2}$ con m racional, para las cuales es bastante sencillo probar (fig. 19.1) que

$$\frac{m}{m+2}\int_{x_0}^{x_1} \frac{dx}{\sqrt{1 + (x/a)^m}} = \text{arc } PP_1 - (P_1 R_1 - PR),$$

donde x_0 y x_1 son las abcisas de R y R_1, y PR y P_1R_1 son las tangentes en P y P_1, respectivamente. Análogamente,

$$\frac{m}{m+2}\int_{z_0}^{z_1} \frac{dz}{\sqrt{1 + (z/a)^m}} = \text{arc } QQ_1 - (Q_1 S_1 - QS).$$

[22] *Acta Erud.*, oct. 1698, 462 y sigs. = *Opera*, 1, 249-253.
[23] *Giornale dei Letterati d'Italia*, vols. 19 y sigs.

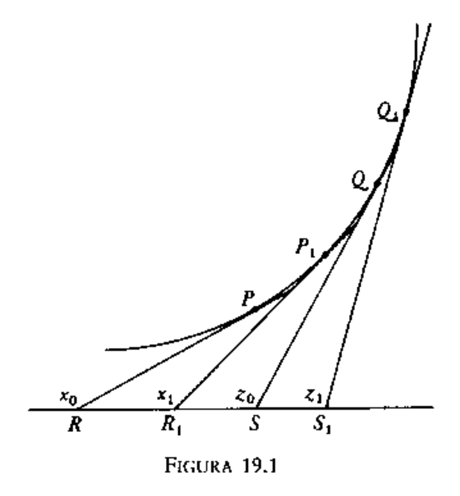


FIGURA 19.1

Por lo tanto, si para alguna relación entre x y z se tiene que

$$\frac{dx}{\sqrt{1 + (x/a)^m}} + \frac{dz}{\sqrt{1 + (z/a)^m}} = 0, \tag{8}$$

entonces la suma de las dos integrales definidas sería 0 y tendríamos

$$\text{arc } QQ_1 - \text{arc } PP_1 = (Q_1S_1 - QS) - (P_1R_1 - PR). \tag{9}$$

Una solución de (8) para $m = 4$ es

$$\frac{x}{a} \cdot \frac{z}{a} = 1. \tag{10}$$

Así pues, con $m = 4$, se tiene que sobre la curva $y = x^3/3a^2$ la diferencia de dos arcos cuyos valores están en la relación (10) se puede expresar como un segmento rectilíneo. Fagnano obtuvo también integrales de (8) para $m = 6$ y $m = 3$.

Fagnano probó además que sobre la elipse, lo mismo que sobre la hipérbola, se pueden encontrar infinitos arcos tales que la diferencia de cada dos de ellos se puede expresar algebraicamente, incluso aunque individualmente los arcos no se puedan rectificar. Así,

en 1716 probó que la diferencia de dos arcos elípticos cualesquiera es algebraica. Analíticamente, partía de

$$\frac{\sqrt{hx^2 + l}}{\sqrt{fx^2 + g}}\,dx + \frac{\sqrt{hz^2 + l}}{\sqrt{fz^2 + g}}\,dz = 0 \qquad (11)$$

o, simplificando,

$$X\,dx + Z\,dz = 0,$$

donde h, l, f, g, x y z satisfacen la condición

$$fhx^2z^2 + flx^2 + flz^2 + gl = 0. \qquad (12)$$

Fagnano probó que

$$\int X\,dx + \int Z\,dz = -\frac{hxz}{\sqrt{-fl}}. \qquad (13)$$

Lo que esto significa geométricamente es que si $2a$ es el eje menor FA de una elipse (fig. 19.2), $CH = x$, $CE = z$, JH es la ordenada en H y GE la ordenada en E, entonces

$$\text{arc } JD + \text{arc } DG = \frac{-hxz}{2a^2} + C. \qquad (14)$$

(Para identificar esto con las integrales, sea p el parámetro *[latus rectum]* de la elipse y sean $p - 2a = h$, $l = 2a^3$, $f = -2a$, $g = 2a^3$. En-

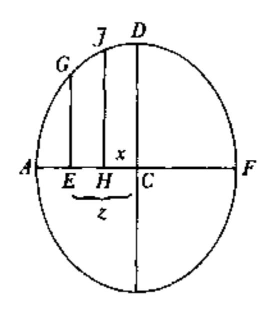


FIGURA 19.2

tonces z es $a\sqrt{2a^3 - 2ax^2} / \sqrt{2a^3 + bx^2}$.) Cuando $x = 0$, arc JD es nulo y el término algebraico de (14) se anula. Por (12), $z = a$, y entonces arc DG se convierte en arc DA; este es el valor de C, con lo que

$$\text{arc } JD + \text{arc } GD = \frac{-bxz}{2a^2} + \text{arc DA},$$

de donde

$$\text{arc } JD - \text{arc } GA = \frac{-bxz}{2a^2}.$$

Un resultado de este trabajo [24], conocido todavía como teorema de Fagnano y obtenido en 1716, establece lo siguiente: sea

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

la elipse de excentricidad e y sean $P(x,y)$ y $P'(x',y')$ (fig. 19.3) cuyos ángulos excéntricos π y ϕ' satisfacen la condición

$$\text{tg } \phi \text{ tg } \phi' = \frac{b}{a}. \tag{15}$$

Entonces el teorema afirma que

$$\text{arc } BP + \text{arc } BP' - \text{arc } BA = e^2 \, xx'/a. \tag{16}$$

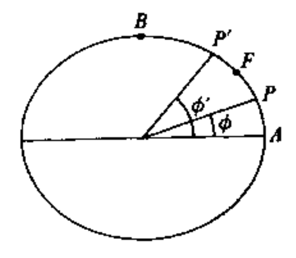


FIGURA 19.3

[24] *Opere*, 2, 287-292.

Los puntos P y P' pueden coincidir, satisfaciendo (15), y para esta posición común F, llamada punto de Fagnano, éste mostró que

$$\text{arc } BF - \text{arc } AF = a - b. \tag{17}$$

A partir de 1714, Fagnano se ocupó también de la rectificación de la lemniscata mediante arcos elípticos e hiperbólicos y en 1717 y 1720 logró integrar otras combinaciones de diferenciales. Así, probó que

$$\frac{dx}{\sqrt{1 - x^4}} = \frac{dy}{\sqrt{1 - y^4}} \tag{18}$$

tiene la integral

$$x = -\sqrt{\frac{1 - y^2}{1 + y^2}}, \tag{19}$$

o sea,

$$x^2 + y^2 + x^2 y^2 = 1. \tag{20}$$

Una manera de enunciar este resultado es: entre dos integrales que expresan arcos de lemniscata (con $a = 1$) existe una relación algebraica, incluso aunque cada integral por separado sea una función trascendente de una nueva clase. Después de esto, Fagnano estableció otras relaciones análogas que le permitieron obtener resultados especiales sobre la lemniscata [25]. Por ejemplo, probó que si

$$\frac{dx}{\sqrt{1 - x^4}} = \frac{2dy}{\sqrt{1 - y^4}}, \tag{21}$$

entonces

$$\frac{\sqrt{1 - y^4}}{y\sqrt{2}} = \frac{\sqrt{1 - x}}{\sqrt{1 - x}}, \tag{22}$$

[25] *Giornale dei Letterati d'Italia*, 30, 1718, 87 y sigs. = *Opere*, 2, 304-313.

y despejando la x,

$$x = \frac{-1 + 2y^2 + y^4}{1 + 2y - y^4}. \quad (23)$$

Mediante diversos resultados de este tipo, Fagnano mostró cómo encontrar los puntos de la lemniscata (o sea, los valores de r en $r^2 = a^2 \cos 2\phi$) que dividen al cuadrante, es decir, el arco CQA en la figura 19.4, en n partes iguales para ciertos valores de n. Probó también cómo, dado un arco CS, se puede hallar el punto I sobre ese arco que lo divide en dos partes iguales; encontró, además, los puntos sobre el arco CQA que, unidos a C, dividen el área comprendida entre CQA y el eje horizontal en dos, tres y cinco partes; y, dadas las cuerdas que dividen dichas área en n partes iguales, determinó las cuerdas que bisecan cada una de esas partes.

Así pues, Fagnano hizo más que dar respuesta a la cuestión de Bernoulli, mostrando que la misma notable propiedad algebraica que caracterizaba las integrales que representaban las funciones trigonométricas y la logarítmica se verificaba para ciertas clases, al menos, de integrales elípticas.

Alrededor de 1750, Euler prestó atención al trabajo de Fagnano sobre la elipse, la hipérbola y la lemniscata y comenzó una serie de investigaciones por su cuenta. En el artículo «Observationes de Compartione Arcuum Curvarum Irrectificabilium» [26], Euler, después de repetir parte del trabajo de Fagnano, mostró cómo dividir el área de un cuadrante de la lemniscata en $n + 1$ partes supuesto que ya está dividida en n partes. Señala luego que su trabajo y el de Fagnano proporcionaban varios resultados útiles sobre integración, y así la

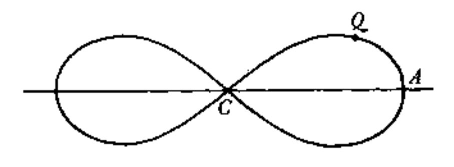


FIGURA 19.4

[26] *Novi Comm. Acad. Sci. Petrop.*, 6, 1756/7, 58-84, pub. 1761 = *Opera*, (1), 20, 80-107.

ecuación (18) tenía, aparte de la integral obvia $x = y$, la integral particular adicional

$$x = -\sqrt{(1-y^2)/(1+y^2)}.$$

En su artículo «De Integratione Aequationis Differentialis $m\,dx/\sqrt{1-x^4} = n\,dy/\sqrt{1-y^4}$» [27], Euler toma los resultados de Fagnano como punto de partida. Las integrales que éste había obtenido para la mayoría de ecuaciones diferenciales que había considerado eran integrales particulares de carácter algebraico; pero la integral completa muy bien podía ser trascendente. Euler se propuso buscar integrales completas en forma algebraica; comenzó con la ecuación (18), pero esperaba obtener la integral completa de

$$\frac{m\,dx}{\sqrt{1-x^4}} = \frac{n\,dy}{\sqrt{1-y^4}}. \tag{24}$$

Aquí, m/n es racional y la ecuación expresa el problema de encontrar dos arcos de lemniscata que están en esa razón. Euler dice que mediante tanteos llegó a la convicción de que (24) posee una integral completa que se puede expresar algebraicamente cuando m/n es racional.

De las investigaciones de Fagnano se seguía que la ecuación (18) se satisfacía por la integral particular (19) o (20). La integral de cada miembro de (18) es un arco de lemniscata con semieje 1 y y abscisa x, y la integración de la ecuación diferencial ordinaria (18) equivale a encontrar dos arcos de la misma longitud. Euler había indicado que $x = y$ era otra integral particular de (18), con lo que la integral completa había de reducirse a esas dos integrales particulares para valores especiales de la constante arbitraria. Guiado por estos hechos, Euler encontró que la integral completa de (18) era

$$x^2 + y^2 + c^2y^2x^2 = c^2 + 2xy\sqrt{1-c^4}, \tag{25}$$

o sea

$$x = \frac{y\sqrt{1-c^4} \pm c\sqrt{1-y^4}}{1+c^2y^2}, \tag{26}$$

[27] *Novi Comm. Acad. Sci. Petrop.*, 6, 1756/7, 35-57, pub. 1761 = *Opera*, (1), 20, 58-79.

donde c es una constante arbitraria. Efectivamente, dada (25), se puede comprobar inmediatamente que es la integral completa de (18).

En el resultado (25) está implícito lo que se conoce a menudo como teorema de adición de Euler para estas integrales elípticas simples. Es claro por simple derivación que

$$\int_0^x \frac{dx}{\sqrt{1-x^4}} = \int_0^y \frac{dx}{\sqrt{1-x^4}} + \int_0^c \frac{dx}{\sqrt{1-x^4}}, \qquad (27)$$

donde c es una constante, es también una integral completa de (18); en consecuencia, x, y, c satisfarán la relación (25) y, así, el teorema de adición afirma que, si se tiene (27) para las integrales elípticas en cuestión, entonces el límite superior x es una función algebraica simétrica, a saber, (26), de los límites superiores y, c, arbitrariamente elegidos, de las otras dos integrales. El teorema de adición es válido para integrales más generales, como veremos.

Utilizando los resultados (25) y (27) es bastante sencillo probar que si

$$\int_0^y \frac{dx}{\sqrt{1-x^4}} = n \int_0^x \frac{dx}{\sqrt{1-x^4}}, \qquad (28)$$

entonces y es una función algebraica de x. Este resultado se conoce como teorema de multiplicación de Euler para la integral elíptica $\int_0^x dx/\sqrt{1-x^4}$. A partir de este resultado se obtiene la integral completa de la ecuación (28); lo importante es que se trata de una ecuación algebraica en x, y y una constante arbitraria c. Euler muestra cómo se puede obtener esa integral completa pero no la da explícitamente.

En el mismo artículo de 1756-67 y en el volumen 7 de la misma revista [28], Euler abordó integrales elípticas más generales; expone el siguiente resultado que, según dice, obtuvo por medio de tanteos. Si se diferencia

$$\alpha + 2\beta(x+y) + \gamma(x^2+y^2) + 2xy + 2\varepsilon xy(x+y) + \xi x^2y^2 = 0, \qquad (29)$$

[28] *Novi Comm. Acad. Sci. Petrop.*, 6, 1758/9, 3-48, pub. 1761 = *Opera*, (1), 20, 153-200.

entonces la ecuación diferencial se puede poner en la forma

$$\frac{dx}{\sqrt{X}} + \frac{dy}{\sqrt{Y}} = 0, \tag{30}$$

donde X e Y son polinomios de cuarto grado de los que cuatro coeficientes (los mismos en X e Y) se pueden expresar en términos de los cinco de (29) con la ayuda de una constante arbitraria. Por lo tanto, (29) es la integral completa de (30) y cuando (30) se particulariza a (18), entonces (29) se convierte en (25). Euler señala que es extraordinario que, aunque la integral de $dx/\sqrt{X}$ no se puede obtener en términos de funciones circulares ni de la logarítmica, la ecuación (30) es satisfecha por una relación algebraica. A continuación, generaliza esos resultados a

$$\frac{m\,dx}{\sqrt{X}} = \frac{n\,dx}{\sqrt{Y}},\ m/n \text{ racional}, \tag{31}$$

donde X e Y son polinomios de cuarto grado con los mismos coeficientes. Lo anterior aparece también en su obra *Institutiones Calculi Integralis* [29], en la cual Euler explica el significado geométrico de tales resultados en términos de las curvas elipse, hipérbola y lemniscata.

A partir de estos resultados, Euler llegó a lo que ahora se conoce como teorema de adición para integrales elípticas de primera especie. Consideremos la integral elíptica

$$\int \frac{dx}{\sqrt{R(x)}}, \tag{32}$$

donde $R(x) = Ax^4 = Bx^3 + Cx^2 + Dx + E$. Entonces, el teorema de adición establece que la ecuación

$$\frac{dx}{\sqrt{R(x)}} = \frac{dy}{\sqrt{R(y)}} \tag{33}$$

es satisfecha por una cierta ecuación algebraica bien definida en x e y tal que y se puede expresar racionalmente en términos de x, el

[29] Volumen 1, sec. 2, cap. 6 = *Opera,* (1), 11, 391-423.

correspondiente valor de $\sqrt{R(x)}$, las constantes arbitrarias x_0 e y_0 y los correspondientes valores de $\sqrt{R(x_0)}$ y $\sqrt{R(y_0)}$. Asimismo, y toma el valor y_0 arbitrariamente prefijado cuando x toma el valor x_0 arbitrariamente dado.

Este resultado conduce a otro teorema que puede ser más ilustrativo. Si la suma o la diferencia de dos integrales elípticas de la forma

$$\int \frac{dx}{\sqrt{R(x)}}, \tag{34}$$

se iguala a una tercera integral de esa forma, y si además el límite inferior de integración y los coeficientes en el radicando son los mismos para las tres integrales, entonces el límite superior de integración de la tercera integral es una función algebraica de los otros dos límites superiores, el límite inferior común y los correspondientes valores de $\sqrt{R(x)}$ en estos tres últimos límites.

Euler fue más allá. Así como el tratamiento de Fagnano de la diferencia de dos arcos de lemniscata lo condujo a la integral elíptica general de primera especie, lo realizado por el mismo Fagnano para la diferencia de dos arcos de elipse (ver (11)) condujo a Euler a un teorema de adición para una segunda clase de integrales [30]. Euler se lamentó de que sus métodos no se pudieran extender a raíces superiores a la raíz cuadrada o a radicandos de grado superior a cuatro y vio, por otro lado, un grave defecto en su trabajo en que no había obtenido sus integrales completas algebraicas por un método general de análisis; así, sus resultados no se podían relacionar de manera natural con otras partes del cálculo infinitesimal.

La obra definitiva sobre integrales elípticas fue realizada por Adrien-Marie Legendre (1752-1833), un profesor de la École Militaire que formó parte de varios comités gubernamentales; más tarde fue examinador de estudiantes en la École Polytechnique. Hasta su muerte en 1833 nunca dejó de trabajar con pasión y regularidad. Su nombre pervive en un gran número de teoremas, muy variados porque abordó las más diversas cuestiones. Sin embargo, no tuvo la originalidad ni la profundidad de Lagrange, Laplace o Monge; su trabajo dio origen a teorías muy importantes, pero sólo después de

---

[30] *Novi Comm. Acad. Sci. Petrop.*, 7, 1758/9, 3-48, pub. 1761 = *Opera*, (1), 20, 153-200, e *Inst. Cal. Integ.*, 1, ¶645 = *Opera*, (1), 11, ¶645.

que fuese asumido por inteligencias más profundas; su nivel se sitúa justo detrás de sus tres contemporáneos arriba citados.

Los teoremas de adición de Euler constituían los principales resultados de la teoría de integrales elípticas cuando Legendre se ocupó del tema en 1786. Durante cuatro décadas fue el único que aportó nuevas investigaciones sobre dichas integrales a la literatura; dedicó dos artículos fundamentales al tema [31], y después escribió los *Exercices de calcul intégral* (3 vols., 1811, 1817, 1826), el *Traité des fonctions elliptiques* [32] (2 vols., 1825-26) y tres suplementos dando cuenta del trabajo de Abel y Jacobi en 1829 y 1832. Los resultados de Euler, como los de Fagnano, estaban ligados a consideraciones geométricas, mientras que Legendre se centró en lo analítico.

El resultado principal de Legendre, que aparece en su *Traité*, consistió en probar que la integral elíptica general

$$\int \frac{P(x)}{\sqrt{R(x)}}\, dx, \tag{35}$$

donde $P(x)$ es una función racional cualquiera de x y $R(x)$ el habitual polinomio general de cuarto grado, se puede reducir a uno de los tres tipos siguientes:

$$\int \frac{dx}{\sqrt{1-x^2}\,\sqrt{1-l^2 x^2}}, \tag{36}$$

$$\int \frac{x^2\, dx}{\sqrt{1-x^2}\,\sqrt{1-l^2 x^2}}, \tag{37}$$

$$\int \frac{dx}{(x-a)\sqrt{1-x^2}\,\sqrt{1-l^2 x^2}}. \tag{38}$$

Legendre denominó estos tres tipos integrales elípticas de primera, segunda y tercera especie, respectivamente. Demostró también que,

[31] *Hist. de l'Acad. des Sci., Paris,* 1786, 616-643 y 664-683.

[32] El uso de la palabra «función» en este texto es engañoso. Legendre estudió las integrales elípticas, en ocasiones con los límites superiores variables, en cuyo caso dichas integrales son, por supuesto, funciones de su límite superior. Pero el término de funciones elípticas se refiere en la actualidad a las funciones introducidas más tarde spor Abel y Jacobi.

por transformaciones ulteriores, esas integrales se pueden reducir a las tres formas siguientes:

$$F(k,\phi)=\int_0^{\phi}\frac{d\phi}{\sqrt{1-k^2\ \text{sen}^2\ \phi}}, \qquad 0<k<1 \quad (39)$$

$$E(k,\phi)=\int_0^{\phi}\sqrt{1-k^2\ \text{sen}^2\ \phi}\quad d\phi, \qquad 0<k<1 \quad (40)$$

$$\pi(n,k,\phi)=\int_0^{\phi}\frac{d\phi}{(1+n\ \text{sen}^2\ \phi)\sqrt{1-k^2\ \text{sen}^2\ \phi}},$$

$$0<k<1 \quad (41)$$

donde n es una constante cualquiera. En estas formas se ve que los valores de las integrales desde $\phi = 0$ hasta $\phi = \pi/2$ son los mismos que desde $\phi = \pi/2$ hasta $\pi = \pi$. La notación Δ (k, ϕ) para la función $\sqrt{1 - k^2 \text{sen}^2 \phi}$ se debe también a Legendre.

Esas formas se pueden convertir, mediante el cambio de variable $x = \text{sen}\ \phi$, en las formas de Jacobi:

$$F(k,x)=\int_0^{x}\frac{dx}{\sqrt{1-x^2}\sqrt{1-k^2x^2}}, \qquad (42)$$

$$\text{E}(k,x)=\int_0^{x}\sqrt{\frac{1-k^2x^2}{1-x^2}}\,dx, \qquad (43)$$

$$n(n,k,x)=\int_0^{x}\frac{dx}{(1+nx^2)\sqrt{(1-x^2)(1-k^2x^2)}}. \quad (44)$$

La cantidad k se denomina módulo de la integral elíptica correspondiente. Si los límites de integración son $\phi = \pi/2$ o $x = 1$, entonces se dice que las integrales son completas; si no, incompletas.

El trabajo de Legendre sobre integrales elípticas tuvo mucho mérito; extrajo numerosas conclusiones, no enunciadas anteriormente, del trabajo de sus predecesores y estructuró matemáticamente dicha materia; sin embargo, no añadió nuevas ideas ni alcanzó la profundidad y penetración de Abel y Jacobi (cap. 27, sec. 6), quienes invirtieron esas integrales, introduciendo así las *funciones* elípticas. Le-

gendre llegó a conocer el trabajo de Abel y Jacobi, a los cuales elogió con mucha humildad y, probablemente, cierta amargura. Al dedicar los suplementos a su trabajo de 1835 a las nuevas ideas de aquéllos, comprendió muy bien que éstas arrojaban a las sombras todo lo que él había hecho en la materia; pasó al lado de uno de los grandes descubrimiento de su época.

5. Otras funciones especiales

Las integrales indefinidas elípticas son funciones trascendentes nuevas. Al irse desarrollando el trabajo analítico del siglo XVIII, se fueron obteniendo más funciones trascendentes, de las que la más importante es la función gamma, surgida de los trabajos sobre dos cuestiones, toería de interpolación y antidiferenciación. El problema de la interpolación había sido considerado por James Stirling (1692-1770), Daniel Bernoulli y Christian Goldbach (1690-1764); fue planteado a Euler y éste anunció su solución en una carta del 13 de octubre de 1729 a Goldbach [33]. En una segunda carta, de 8 de enero de 1730, se planteó el problema de la integración [34]. En 1731, Euler publicó resultados sobre ambos problemas en un artículo, «De Progressionibus...» [35].

El problema de la interpolación consistía en dar sentido a $n!$ para valores no enteros de n. Euler indicó que

$$n! = \left[\left(\frac{2}{1}\right)^n \frac{1}{n+1}\right]\left[\left(\frac{3}{2}\right)^n \frac{2}{n+2}\right]\left[\left(\frac{4}{3}\right)^n \frac{3}{n+3}\right]\cdots \qquad (45)$$

$$= \prod_{k=1}^{\infty}\left(\frac{k+1}{k}\right)^n \frac{k}{k+n}.$$

La ecuación resulta formalmente correcta si se simplifican los factores comunes en ese producto infinito. Pero esta expresión analítica de $n!$, a diferencia de la definición original $n(n-1)\cdots 2\cdot 1$, tiene sentido para todo n excepto para los negativos. Euler observó que

[33] Fuss, *Correspondance*, 1, 3-7.

[34] Fuss, *Correspondance*, 1, 11-18.

[35] *Comm. Acad. Sci. Petrop.*, 5, 1730/1, 36-57, pub. 1738 = *Opera*, (1), 24, 1-24.

para $n = 1/2$ el segundo miembro da, después de algunos cálculos, el producto infinito de Wallis

$$\frac{\pi}{2} = \left(\frac{2 \cdot 2}{1 \cdot 3}\right)\left(\frac{4 \cdot 4}{3 \cdot 5}\right)\left(\frac{6 \cdot 6}{5 \cdot 7}\right)\left(\frac{8 \cdot 8}{7 \cdot 9}\right)\cdots \tag{46}$$

En la notación $\Gamma(n+1) = n!$, introducida más tarde por Legendre, Euler demostró también que $\Gamma(n+1) = n\Gamma(n)$, de donde obtuvo $\Gamma(3/2)$, $\Gamma(5/2)$, y así sucesivamente.

Euler podría haber utilizado (45) como su generalización del concepto de factorial y, de hecho, hoy se introduce a menudo en la forma equivalente, también dada por Euler,

$$\lim_{m\to\infty} \frac{m!\,(m+1)^n}{(n+1)(n+2)\cdots(n+m)}. \tag{47}$$

Pero la relación con el resultado de Wallis llevó a Euler a estudiar la integral, ya considerada por Wallis,

$$\int_0^1 x^e(1-x)^n\,dx, \tag{48}$$

en donde e y n son para Euler arbitrarios. Euler evaluó esta integral desarrollando $(1-x)^n$ por el teorema del binomio, obteniendo

$$\int_0^1 x^e(1-x)^n\,dx$$
$$= \frac{1}{e+1} - \frac{n}{1(e+2)} + \frac{n(n-1)}{1 \cdot 2(e+3)} - \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3\,(e+4)} + \cdots. \tag{49}$$

Para $n = 0, 1, 2, 3, \ldots$ las sumas del segundo miembro son, respectivamente,

$$\frac{1}{e+1}, \frac{1}{(e+1)(e+2)}, \frac{1 \cdot 2}{(e+1)(e+2)(e+3)},$$
$$\frac{1 \cdot 2 \cdot 3}{(e+1)(e+2)(e+3)(e+4)}, \cdots. \tag{50}$$

Así pues, Euler obtuvo para n entero positivo que

$$\int_0^1 x^e (1-x)^n \, dx$$

$$= \frac{n!}{(e+1)(e+2)\cdots(e+n+1)}. \qquad (51)$$

A continuación, Euler buscó una expresión para $n!$ con n arbitrario; mediante una serie de transformaciones que hoy no nos serían totalmente aceptables, Euler llegó a

$$n! = \int_0^1 (-\log x)^n \, dx. \qquad (52)$$

Esta integral tiene sentido para casi todo n y recibe el nombre de segunda integral euleriana o, como la denominó Legendre más tarde, función gamma, denotándose por $\Gamma(n+1)$. [Gauss escribió $\pi(n) = \Gamma(n+1.]$ Más tarde, en 1781 (pub. 1794), Euler dio la forma moderna, que se obtiene de (52) haciendo $t = -\log x$,

$$\Gamma(n+1) = \int_0^\infty x^n e^{-x} \, dx. \qquad (53)$$

La integral (48), a la que Legendre denominó primera integral euleriana, quedó normalizada como la función beta,

$$B(m, n) = \int_0^\infty x^{m-1} (1-x)^{n-1} \, dx. \qquad (54)$$

Euler descubrió la relación existente entre ambas integrales [36], a saber,

$$B(m, n) = \frac{(m)\Gamma(n)}{\Gamma(m+n)}.$$

En sus *Exercices de calcul intégral*, Legendre llevó a cabo un profundo estudio de las integrales eulerianas, llegando a la fórmula de duplicación

$$\Gamma(2x) = (2\pi)^{-1/2} \, 2^{2x-(1/2)} \, \Gamma(x) \, \Gamma\left(x + \frac{1}{2}\right). \qquad (55)$$

[36] *Novi Comm. Acad. Sci. Petrop.*, 16, 1771, 91-139, pub. 1772 = *Opera*, (1), 17, 316-357.

Gauss estudió la función gamma en su trabajo sobre la función hipergeométrica [37] y extendió el resultado de Legendre a la denominada fórmula de multiplicación:

$$\Gamma(nx) = (2\pi)^{(1-n)/2}\, n^{nx-(1/2)}\, \Gamma(x)\, \Gamma\left(x+\frac{1}{n}\right) \Gamma\left(x+\frac{2}{n}\right) \cdots \Gamma\left(x+\frac{n-1}{n}\right). \qquad (56)$$

6. Funciones de varias variables

El desarrollo del cálculo infinitesimal de funciones de dos y tres variables se inició a comienzos de siglo. Señalaremos solamente algunos detalles.

Aunque Newton obtuvo a partir de ecuaciones polinomiales en x e y, es decir, $f(x,y)=0$, expresiones que hoy obtenemos por derivación parcial de f respecto a x e y, su trabajo no fue publicado. Jacques Bernoulli también utilizó derivadas parciales en su trabajo sobre problemas isoperimétricos, como también hizo Nicolaus Bernoulli (1687-1749) en un artículo de las *Acta Eruditorum* de 1720 sobre trayectorias ortogonales. Sin embargo, fueron Alexis Fontaine des Bertins (1705-71), Euler, Clairaut y D'Alembert quienes crearon la teoría de derivadas parciales.

Al principio, la diferencia entre una derivada parcial y una ordinaria no fue reconocida explícitamente, y se utilizaba el mismo símbolo d para ambas. En el caso de funciones de varias variables independientes, era el significado el que indicaba la derivada de que se trataba, correspondiente a cambios en una única variable.

La condición para que $dz = p\,dx + q\,dy$, donde p y q son funciones de x e y, sea una diferencial exacta, es decir, proviniente de $z=f(x,y)$ al formar la diferencial $dz = (\partial f/\partial x)\,dx + (\partial f/\partial y)\,dy$, fue obtenida por Clairaut [38]. Su resultado fue que $p\,dx + q\,dy$ es una diferencial exacta si y sólo si $\partial p/\partial y = \partial q/\partial x$.

El principal impulso para trabajar con derivadas de funciones de dos o más variables vino de los primeros trabajos en ecuaciones en

[37] *Com. Soc. Gott.*, II, 1813 = *Werke*, 3, 123-162, p. 149 en parte.
[38] *Mém. de l'Acad. des Sci., Paris*, 1739, 425-436, y 1740, 293-323.

derivadas parciales. Así, Euler desarrolló un cálculo de derivadas parciales en una serie de artículos dedicados a problemas de hidrodrinámica. En un artículo de 1734 [39] demuestra que si $z = f(x, y)$, entonces

$$\frac{\partial^2 z}{\partial x\, \partial y} = \frac{\partial^2 z}{\partial y\, \partial x} .$$

En otros artículos escritos de 1748 a 1766 trata de los cambios de variables, inversión de derivadas parciales y de los determinantes funcionales. D'Alembert, en trabajos realizados en 1744 y 1745 sobre dinámica, extendió el cálculo de derivadas parciales.

En el trabajo de Newton, en los *Principia*, sobre la atracción gravitatoria ejercida por esferas y casquetes esféricos sobre partículas, están ciertamente implicadas integrales múltiples, pero Newton utilizó argumentos geométricos. En el siglo XVIII el trabajo de Newton fue reformulado analíticamente y ampliado. Aparecen integrales múltiples en la primera mitad del siglo, siendo utilizadas para denotar la solución de $\partial^2 z/\partial x\, \partial y = f(x, y)$ y también, por ejemplo, para determinar la atracción gravitatoria ejercida por una lámina sobre partículas. Así, la atracción de una lámina elíptica de grosor δc sobre un punto situado directamente sobre el centro a una distancia de c unidades es igual a una constante por la integral

$$\delta c \iint \frac{c\, dx\, dy}{(c^2 + x^2 + y^2)^{3/2}}$$

evaluada sobre la elipse $(x^2/a^2) + (y^2/b^2) = 1$. Esta integral fue calculada por Euler en 1738 por integración iterada [40], integrando respecto a y y desarrollando en serie respecto a x el nuevo integrando.

Para 1770, Euler tenía una idea clara de la integral doble definida sobre un dominio acotado limitado por arcos y dio un procedimiento para calcular tales integrales mediante integración iterada [41]. Lagrange, en su trabajo sobre la atracción ejercida por elipsoides de revolución [42], expresó dicha atracción como una integral triple; en-

---

[39] *Comm. Acad. Sci. Petrop.*, 7, 1734/5, 174-193, pub. 1740 = *Opera*, (1), 22, 36-56.

[40] *Comm. Acad. Sci. Petrop.*, 10, 1738, 102-115, pub. 1747 = *Opera*, (2), 6, 175-188.

[41] *Novi Comm. Acad. Sci. Petrop.*, 14, 1769, 72-103, pub. 1770 = *Opera*, (1), 17, 289-315.

[42] *Nouv. Mém. de l'Acad. de Berlin*, 1773, 121-148, pub. 1775 = *Œuvres*, 3, 619-658.

contrando difícil realizar el cálculo en coordenadas rectangulares, efectuó un cambio a coordenadas cilíndricas, a saber,

$$x = a + r \text{ sen } \phi \cos \theta$$
$$y = b + r \text{ sen } \phi \text{ sen } \theta$$
$$z = c + r \cos \phi,$$

donde a, b y c son las coordenadas del nuevo origen, θ es la longitud, ϕ la colatitud y $0 \leq \phi \leq \pi$, $0 \leq \theta \leq 2\pi$. Lo esencial en la transformación de la integral es reemplazar $dx\,dy\,dz$ por r^2 sen $\theta\,d\theta\,d\phi\,dr$. Comenzó así Lagrange el tema de los cambios de variables en las integrales múltiples; de hecho, desarrolló el método general, aunque no muy claramente. También Laplace dio el cambio a coordenadas esféricas casi simultáneamente [43].

7. Los intentos de proporcionar rigor al cálculo infinitesimal

El desarrollo de los conceptos y las técnicas del cálculo infinitesimal fue acompañado de esfuerzos para dotarlo de los fundamentos de que carecía. Los libros sobre la materia que apareiceron después de los intentos fallidos de Newton y Leibniz para explicar los conceptos y justificar los procedimientos empleados intentaron aclarar la confusión existente, pero en realidad la aumentaron.

La manera de enfocar el cálculo infinitesimal de Newton era potencialmente más fácil de rigorizar que la de Leibniz, aunque la metodología de éste era más fluida y más práctica en las aplicaciones. Los ingleses pensaron que podrían conseguir el rigor en ambos enfoques intentando ligarlos a la geometría de Euclides, pero confundieron los momentos de Newton (sus incrementos indivisibles) con sus fluxiones, las cuales se refieren a variables continuas. Los continentales, siguiendo a Leibniz, trabajaron con diferenciales e intentaron dotar de rigor a este concepto. Las diferenciales se consideraban bien como infinitesimales, es decir, cantidades no nulas pero tampoco de ningún tamaño finito, o, a veces, como cero.

Brook Taylor (1685-1731), que fue secretario de la Royal Society de 1714 a 1718, intentó, en sus *Methodus Incrementorum Directa et Inversa* (1715), clarificar las ideas del cálculo infinitesimal, aunque

---

[43] *Mém. des sav. étrangers,* 1772, 536-544, pub. 1776 = *Œuvres,* 8, 369-477.

limitándose a funciones algebraicas y ecuaciones diferenciales algebraicas; pensó que podía considerar siempre incrementos finitos pero fue impreciso acerca de la transición de éstos a fluxiones. La exposición de Taylor, basada en lo que podríamos llamar diferencias finitas, no logró muchos partidarios a causa de su naturaleza aritmética cuando los británicos estaban intentando relacionar el cálculo inifinitesimal con la geometría o con la noción física de velocidad.

Se puede también apreciar algo de la oscuridad y del fracaso de los esfuerzos del siglo XVIII en la obra de Thomas Simpson (1710-61) *A New Treatise on Fluxions* (1737). Después de algunas definiciones preliminares, define así una fluxión: «La magnitud en la que cualquier cantidad fluente sería uniformemente incrementada en una porción dada de tiempo con la celeridad generadora en una posición o instante dados (si permaneciese invariable desde entonces) es la fluxión de dicha cantidad en esa posición o instante.» En nuestro lenguaje, Simpson está definiendo la derivada diciendo que es $(dy/dx)\,\Delta t$. Algunos autores se dieron por vencidos. El matemático francés Michel Rolle señalaba en una ocasión que el cálculo infinitesimal era una colección de falacias ingeniosas.

El siglo XVIII asistió también a nuevos ataques al cálculo infinitesimal. El más duro fue el realizado por el obispo George Berkeley (1685-1753), quien temía la creciente amenaza planteada a la religión por el mecanicismo y el determinismo. En 1734 publicó *The Analyst, Or A Discourse Addressed to an Infidel Mathematician. Wherein It is examied whether the Objet, Principles and inferences of the modern Analysis are more distinctly conceived, or more evidently deduced, than Religious Mysteries and Points of Faith. «First cast out the beam out of thine Eye; and then shalt thou see clearly to cast out the mote of thy brother's Eye.»* (El «infiel» era Edmond Halley.) [44]

Berkeley señaló con razón que los matemáticos estaban procediendo más bien inductiva que deductivamente y que no daban la lógica o las razones de sus pasos. Criticó muchos de los argumentos de Newton; por ejemplo, en el tardío *De Quadratura*, en el que dice que había evitado lo infinitamente pequeño, da a x un incremento denotado por o, desarrolla $(x + o)^n$, resta x^n, divide por o para hallar la razón de los incrementos de x^n y x y a continuación desprecia los términos que contienen o obteniendo así la fluxión de x^n.

[44] George Berkeley, *The Works*, G. Bell and Sons, 1898, vol. 3, 1-51.

Berkeley dice que Newton da primero a x un incremento pero que después lo hace cero; esto, dice, es un desafío a la ley de contradicción y la fluxión obtenida es, en realidad, 0/0. Berkeley atacó también el método de diferenciales según lo presentaban L'Hospital y otros en el continente; la razón de las diferenciales, decía, determinaría la secante y no la tangente; el error se anula al despreciar diferenciales de orden superior y así «en virtud de un doble error, aunque no a una ciencia, se llega a pesar de todo a la verdad», gracias a que los errores se compensaban uno al otro. Criticó también la segunda diferencial $d(dx)$ por ser la derivación de una cantidad, dx, ya por sí misma poco menos que imperceptible; escribe: «En cualquier otra ciencia los hombres demuestran las conclusiones a partir de los principios y no los principios a partir de las conclusiones.»

En lo que atañe a la derivada contemplada como la razón de los incrementos evanescentes en y y x, o sea, dy y dx, éstos no eran «ni cantidades finitas, ni cantidades infinitamente pequeñas ni siquiera nada». Estas razones de cambio no eran sino «los espectros de las cantidades difuntas. Ciertamente... quien pueda digerir una segunda o tercera fluxión... no ha de hacer, creo yo, remilgos a ningún argumento de la teología». Concluía Berkeley que los principios de las fluxiones no eran más claros que los del cristianismo y rechazó que el objeto, los principios y las inferencias del análisis moderno estuviesen concebidos más claramente ni deducidos más sólidamente que los misterios religiosos o lo argumentos de la fe.

El *Analyst* fue replicado por James Jurin (1684-1750), el cual publicó en 1734 *Geometry, No Friend to Infidelity*, en donde mantenía que las fluxiones eran claras para aquellos versados en geometría. Jurin intentó sin éxito explicar los momentos y las fluxiones de Newton; por ejemplo, definía el límite de una cantidad variable como «cierta cantidad determinada a la que la cantidad variable se supone que se aproxima continuamente» estando más cerca de ella que cualquier diferencia dada, pero a continuación añadía «sin sobrepasarla nunca»; aplicó esta definición a una razón variable (el cociente incremental). La demoledora respuesta de Berkeley, titulada *A Defense of Freethinking in Mathematics* (1735) [45], afirmaba que Jurin estaba tratando de defender lo que no comprendía; Jurin replicó pero no clarificó la cuestión.

Entró entonces en la refriega Benjamin Robins (1707-51) con

---

[45] George Berkeley, *The Works*, G. Bell and Sons, 1898, vol. 3, 53-89.

artículos y un libro, *A Discourse Concerning the Nature and Certainty of Sir Isaac Newton's Method of Fluxions and of Prime and Ultimate Ratios* (1735). Robins dejó de lado los momentos del primer artículo de Newton, subrayando, sin embargo, la importancia de las fluxiones y de las razones primeras y últimas; definía un límite así: «Definimos un límite como una magnitud última, a la cual una magnitud variable se puede acercar con cualquier grado de aproximación, aunque pueda no llegar a hacerse nunca totalmente igual a ella.» En su opinión, las fluxiones eran la idea correcta, mientras que las razones primeras y últimas representaban únicamente una explicación; dijo también que el método de fluxiones se establecía sin recurrir a límites, a pesar de que él dio explicaciones en términos de variables que se aproximan a un límite, y rechazó los infinitesimales.

Para responder a Berkeley, Colin Maclaurin (1698-1746), en su *Treatise of Fluxions* (1742), intentó dotar de rigor al cálculo infinitesimal; fue un esfuerzo encomiable pero fallido. Lo mismo que Newton, Maclaurin amaba la geometría, y por ello trató de fundamentar la doctrina de las fluxiones en la geometría de los griegos y el método de exhausción, en particular, tal como fue utilizado por Arquímedes, esperando de este modo evitar el concepto de límite; su logro fue utilizar tan hábilmente la geometría que persuadió a otros a hacer lo mismo y abandonar el análisis.

Los matemáticos continentales se fiaban más de las manipulaciones formales de expresiones algebraicas que de la geometría. El representante más importante de este enfoque fue Euler, quien rechazaba la geometría como base para el cálculo infinitesimal y trató de trabajar con funciones de una manera puramente formal, es decir, razonando a partir de su representación algebraica (analítica).

Euler rechazó el concepto de infinitesimal, una cantidad menor que cualquier cantidad fijada y sin embargo no nula. En sus *Institutiones* de 1755 sostenía que [46]:

> No hay duda de que cualquier cantidad puede disminuirse hasta tal punto que se anule completamente y desaparezca. Pero una cantidad infinitamente pequeña no es otra cosa que una cantidad evanescente y por tanto ella misma ha de ser igual a 0. Ello está también en armonía con esa definición de cosas infinitamente pequeñas según la cual se dice que son menores que cualquier cantidad fijada; ciertamente, debería ser nula, pues si no fuese

---

[46] *Opera*, (1), 10, 69.

igual a 0 se le podría asignar una cantidad igual, lo que es contrario a la hipótesis.

Puesto que Euler desterró las diferenciales tenía que explicar cómo dy/dx, que para él era 0/0, podía ser igual a un número bien definido. Lo hizo de la siguiente manera: dado que para cualquier número n se tiene que $n \cdot 0 = 0$, entonces $n = 0/0$; la derivada es simplemente un método útil de determinar 0/0; para justificar el despreciar $(dx)^2$ en presencia de dx, Euler afirma que $(dx)^2$ se anula antes de que lo haga dx, de modo que, por ejemplo, la razón de $dx + (dx)^2$ a dx es igual a 1. Aceptó ∞ como número, por ejemplo, como la suma $1 + 2 + \cdots$, y también distinguió órdenes de ∞. Así, $a/0 = \infty$, pero $a/(dx)^2$ es un infinito de segundo orden, y así sucesivamente. Procede entonces Euler a obtener la derivada de $y = x^2$ como sigue: da a x el incremento ω; el correspondiente incremento de y es $\eta = 2x\omega + \omega^2$ y la razón η/ω vale $2x + \omega$; dice entonces que esta razón se aproxima tanto más a $2x$ cuanto más pequeño se toma ω, pero recalca que estas diferenciales η y ω son absolutamente cero y que no se puede deducir de ellas otra cosa que su razón mutua, la cual se reduce finalmente a una cantidad finita. Así, Euler acepta incondicionalmente que existen cantidades que son absolutamente cero pero cuyas razones son números finitos. Hay más «razonamientos» de esta naturaleza en el capítulo 3 de las *Institutiones*, donde alienta al lector señalando que no hay tanto misterio oculto en la derivada como se piensa, lo que provoca sospechas sobre el cálculo infinitesimal en el espíritu de tantos.

Como ejemplo adicional de los razonamientos de Euler, consideremos su derivación de la diferencial de $y = \log x$, en la sección 180 de sus *Institutiones* (1775). Reemplazando x por $x + dx$ se tiene

$$dy = \log\,(x + dx) - \log x = \log\left(1 + \frac{dx}{x}\right).$$

Apela aquí a un resultado del capítulo 7 del volumen 1 de su *Introductio* (1748).

$$\log_e\,(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots. \qquad (57)$$

Reemplazando x por dx/x se obtiene

$$dy = \frac{dx}{x} - \frac{dx^2}{2x^2} + \frac{dx^3}{3x^3} - \cdots.$$

Como todos los términos después del primero son evanescentes, tenemos

$$d(\log x) = \frac{dx}{x}.$$

Hemos de tener presente que los textos de Euler eran lo habitual en su tiempo. A lo que contribuyó Euler con su enfoque formalista fue a liberar el cálculo infinitesimal de la geometría y a basarlo en la aritmética y el álgebra; este paso sirvió, cuando menos, para preparar la justificación final de aquél sobre la base del sistema de los números reales.

Lagrange, en un artículo de 1772 [47] y en su *Théorie des fonctions analytiques* [48], llevó a cabo el intento más ambicioso de reconstruir los fundamentos del cálculo infinitesimal. El subtítulo de su libro revela su desvarío; dice así: «Conteniendo los principales teoremas del cálculo diferencial sin hacer uso de lo infinitamente pequeño, ni de cantidades evanescentes ni de límites o fluxiones, y reducido al arte del análisis algebraico de cantidades finitas.»

Lagrange critica el enfoque de Newton señalando que, en lo que se refiere a la razón límite del arco a la cuerda, Newton considera iguales arco y cuerda no antes o después de desvanecerse sino *cuando* se desvanecen. Como señala correctamente Lagrange, «Este método tiene el inconveniente de considerar cantidades en el estado en que, por así decirlo, cesan de ser cantidades; pues aunque siempre podemos concebir correctamente las razones de dos cantidades mientras ellas permanecen finitas, la mente no se hace una idea clara y precisa de esa razón cuando sus términos se convierten, ambos al mismo tiempo, en nada.» El *Treatise of Fluxions* de Maclaurin muestra, dice Lagrange, lo difícil que es justificar el método de fluxiones. Tampoco le satisfacen los ceros pequeños (infinitesimales) de Leibniz y Bernoulli ni los ceros absolutos de Euler, todos los cuales,

[47] *Nouv. Mém. de l'Acad. de Berlin*, 1772, pub. 1774 = *Œuvres*, 3, 441-476.

[48] 1797; 2.ª ed., 1813 = *Œuvres*, 9.

«aunque en realidad correctos no son lo suficientemente claros como para servir de fundamento a una ciencia cuya certeza debe reposar en su propia evidencia».

Lagrange quiso dotar al cálculo infinitesimal del rigor de las demostraciones de los antiguos y propuso lograr esto reduciéndolo al álgebra, la cual, como señalamos más arriba, incluía la series como extensiones de polinomios. En efecto, la teoría de funciones es para Lagrange la parte del álgebra que se refiere a las derivadas de funciones; concretamente, Lagrange propuso utilizar series de potencias, señalando con discreta modestia su extrañeza de que este método no se lo hubiese ocurrido a Newton. Se propone entonces hacer uso del hecho de que toda función $f(x)$ se pude expresar en la forma:

$$f(x+h) = f(x) + ph + qh^2 + rh^3 + sh^4 + \cdots, \qquad (58)$$

en donde los coeficientes p, q, r, $\cdots$ contienen x pero son independientes de h; pero quiere estar seguro antes de continuar de que tal desarrollo en serie de potencias es siempre posible. Desde luego, afirma, esto se sabe por numerosos ejemplos familiares, pero concede que hay casos excepcionales; los que Lagrange tiene *in mente* son aquellos en los que alguna derivada de $f(x)$ se hace infinita y aquellos en los que la función y sus derivadas se hacen infinitas; pero estas excepciones sólo ocurren en puntos aislados y Lagrange no las tiene en cuenta; sin mayores miramientos hace frente a una segunda dificultad: tanto Lagrange como Euler aceptaban sin reservas que era perfectamente posible un desarrollo en serie conteniendo potencias enteras y fraccionarias de h pero Lagrange quería eliminar la necesidad de las potencias fraccionarias; éstas surgen, pensaba Lagrange, solamente si $f(x)$ contiene radicales, con lo que también las descarta como casos excepcionales. Deja así todo listo para seguir adelante con (58).

Mediante un argumento un tanto complicado pero puramente formal, Lagrange concluye que podemos obtener $2q$ de p del mismo modo que obtenemos p de $f(x)$, y que una conclusión análoga se tiene también para los demás coeficientes r, s, ... de (58). De aquí, si denotamos p por $f'(x)$ y designamos por $f''(x)$ una función derivada de $f'(x)$ como $f'(x)$ se deriva de $f(x)$, entonces

$$p = f'(x), \qquad q = \frac{1}{2!} f''(x), \qquad r = \frac{1}{3!} f'''(x), \ldots,$$

de donde (58) da

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\cdots .$$

Lagrange concluye entonces que la última «expresión tiene la ventaja de mostrar cómo los términos de la serie dependen uno de otro, y especialmente cómo cuando se sabe formar la primera función derivada, se pueden formar todas las funciones derivadas que intervienen en la serie». Un poco más adelante añade que: «Para quien conoce los rudimentos del cálculo diferencial es claro que estas funciones derivadas coinciden con dy/dx, d^2y/dx^2, $\cdots$.»

A Lagrange le resta todavía mostrar cómo deriva p o $f'(x)$ de $f(x)$. Para ello, utiliza (58) despreciando todos los términos después del segundo. Así pues, $f(x+h)-f(x)=ph$, divide por h y concluye que $p=f'(x)$.

En realidad, la suposición de Lagrange de que una función se puede desarrollar en serie de potencias es uno de los puntos débiles de su esquema. El criterio, hoy conocido, para que tal desarrollo sea posible requiere la existencia de derivadas, y esto es lo que Lagrange pretendía evitar. Sus argumentos para justificar las series de potencias sólo sirvieron para añadir confusión acerca de qué funciones admitían tal desarrollo; incluso cuando éste es posible, Lagrange muestra cómo calcular los coeficientes sólo si conocemos el primero, es decir, $f'(x)$; y en cuanto a éste, utiliza los mismos toscos argumentos de sus predecesores. Finalmente, la cuestión de la convergencia de la serie (58), en rigor no la plantea; prueba que para h suficientemente pequeño el último término que se conserva es mayor que lo que se desprecia y también da en este libro la forma de Lagrange del resto en un desarrollo de Taylor (cap. 20, sec. 7), pero esto no desempeña ningún papel en los argumentos de más arriba. A pesar de todos estos puntos débiles, el enfoque de Lagrange del cálculo infinitesimal gozó de gran aceptación durante bastante tiempo, siendo más tarde abandonado.

Lagrange creyó que había prescindido del concepto de límite. Reconocía [49] que el cálculo infinitesimal se podía fundamentar sobre una teoría de límites, pero afirmó que la clase de metafísica que era necesario emplear era ajena al espíritu del análisis. A pesar de las

[49] *Œuvres*, 1, 325.

insuficiencias de su método, contribuyó, como lo hizo Euler, a separar la fundamentación del análisis de la geometría y la mecánica; en esto, su influencia fue decisiva. Aunque esta separación no es pedagógicamente deseable, ya que impide la comprensión intuitiva, dejó claro que, desde el punto de vista lógico, el análisis debía desarrollarse por sus propios medios.

Hacia finales de siglo, el matemático, soldado y administrador Lazare N. M. Carnot (1753-1823) escribió una obra popular, muy vendida, *Réflexions sur la métaphysique du calcul inifnitésimal* (1797), en la que pretendió dotar de precisión al cálculo infinitesimal; intentó probar que el fundamento lógico estribaba en el método exhaustivo y que todas las maneras de tratar la materia no eran sino simplificaciones o atajos cuya lógica podría ser justificada fundamentándose sobre dicho método. Después de muchas reflexiones acabó concluyendo, con Berkeley, que los errores en los razonamientos habituales del cálculo infinitesimal se compensaban unos con otros.

Entre la multitud de esfuerzos para rigorizar el cálculo infinitesimal, unos pocos de ellos fueron por el buen camino. Los más notables de éstos fueron los de D'Alembert y, antes, Wallis. D'Alembert pensaba que Newton había tenido la idea correcta y que él simplemente explicaba lo que había querido decir Newton. En su artículo «Différentiel» de la célebre *Encyclopédie ou Dictionnaire Raisonné des Sciences, des Arts et des Métiers* (1751-80), afirma: «Newton nunca contempló el cálculo diferencial como un cálculo de infinitesimales, sino como un método de razones primeras y últimas, es decir, un método para encontrar el límite de estas razones.» Pero D'Alembert definía una diferencial como «una cantidad infinitamente pequeña o al menos más pequeña que cualquier magnitud fijada». Pensaba que el cálculo de Leibniz podía edificarse sobre tres reglas de diferenciales; sin embargo, era partidario de la derivada en tanto que límite; en su objetivo de utilizar límites afirmó, como Euler, que 0/0 podía valer cualquier cantidad que se quisiera.

En otro artículo, «Limite», afirma: «La teoría de límites es la verdadera metafísica del cálculo... No es nunca cuestión de cantidades infinitesimales en el cálculo diferencial: es únicamente una cuestión de límites de cantidades finitas. Así pues, la metafísica de las cantidades infinitas y las infinitamente pequeñas, mayores o menores que otra, es completamente inútil en el cálculo diferencial.» Los infinitesimales eran simplemente una manera de hablar que evitaba las descripciones más extensas en términos de límites. De hecho,

D'Alembert dio una buena aproximación a la definición correcta de límite en términos de una cantidad variable que se aproxima a una cantidad fija con un error menor que cualquier cantidad fijada, aunque aquí también él dice que la variable nunca alcanza el límite. Con todo, D'Alembert no llevó a cabo una exposición formal del cálculo infinitesimal que incorporase y utilizase sus, en esencia, correctas opiniones. Fue también impreciso en cierto número de cuestiones; por ejemplo, definió la tangente a una curva como el límite de la secante cuando los dos puntos de intersección se hacen uno. Esta imprecisión, especialmente en su enunciado de la noción de límite, originó un debate sobre la cuestión de si una variable puede alcanzar su límite. Al no existir una presentación explícita correcta, D'Alembert aconsejaba a los estudiantes de cálculo infinitesimal, «Persistid y os llegará la fe».

Sylvestre-François Lacroix (1765-1843), en la segunda edición (1810-19) de su *Traité du calcul différentiel et du calcul intégral*, vio más explícitamente que la razón de dos cantidades, cada una de las cuales se aproxima a 0, puede aproximarse a un número bien definido al que tiene como límite; considera la razón $ax/(ax+x^2)$ y observa que es la misma que $a/(a+x)$, y que ésta se aproxima a 1 cuando x se aproxima a 0. Además, señala que 1 es el límite incluso cuando x se acerca a 0 con valores negativos; sin embargo, también habla de la razón de los límites cuando éstos valen 0 e incluso utiliza el símbolo 0/0. Introdujo la diferencial dy de una función $y=f(x)$ en términos de la derivada; así, si $y=ax^3$, $dy=3ax^2\,dx$. Utilizó al principio el término de «coeficiente diferencial» para la derivada; así, $3ax^2$ sería el coeficiente diferencial.

Casi todos los matemáticos del siglo XVIII realizaron algún esfuerzo o al menos se pronunciaron acerca de la lógica del cálculo infinitesimal, pero aunque uno o dos de ellos estaban en el buen camino todos los esfuerzos resultaron fallidos. La distinción entre un número muy grande y un «número» infinito difícilmente se hacía; si un teorema era cierto para todo n parecía claro que también lo era para n infinito. De manera análoga, un cociente incremental se reemplazaba por la derivada y una suma de un número finito de términos difícilmente se distinguía de una integral; los matemáticos pasaban de una a otra con toda libertad. En 1755, en su *Institutiones*, Euler distinguía entre el incremento de una función y la diferencial de esa función y entre la sumación y la integral, pero estas distinciones no eran proseguidas. Todos esos esfuerzos podrían resumirse

en la descripción de Voltaire del cálculo infinitesimal como «el arte de numerar y medir exactamente una cosa cuya existencia no puede ser concebida».

A la vista de la ausencia casi total de alguna clase de fundamento, ¿cómo procedían los matemáticos para manipular tal diversidad de funciones? Aparte de su gran confianza en los significados físico e intuitivo, tenían *in mente* un modelo —las funciones algebraicas más simples, como los polinomios y las funciones racionales—; trasladaban a todas las funciones las propiedades que descubrían en estas funciones concretas, explícitas: continuidad, existencia de infinitos y discontinuidades aisladas, desarrollo en serie de potencias, existencia de derivadas e integrales. Pero cuando se vieron obligados, en buena medida a causa de los trabajos sobre la cuerda vibrante, a ampliar el concepto de función, como lo expresó Euler, a cualquier curva de trazo libre (las funciones mixtas, o irregulares, o discontinuas de Euler) ya no les fue posible por más tiempo utilizar como guía las funciones más simples. Y cuando la función logarítmica hubo de extenderse a los números negativos y a los complejos, procedieron ya sin ninguna base fiable en absoluto; esta es la razón por la que eran comunes las discusiones sobre estas materias. La rigorización del cálculo infinitesimal no fue alcanzada hasta el siglo XIX.

Bibliografía

Bernoulli, Jacques: *Opera*, 2 vols., 1744, reimpresos por Birkhaüser, 1968.

Bernoulli, Jean: *Opera Omnia*, 4 vols., 1742, reimpresos por Georg Olms, 1968.

Boyer, Carl B.: *The concepts of the Calculus*, Dover (reimpresión), 1949, cap. 4.

Brill, A. y M. Nöther: «Die Entwicklung der Theorie der algebraischen Funktionen in älterer und neuerer Zeit», *Jahres. der Deut. Math.-Verein*, 3, 1892/3, 107-566.

Cajori, Florian: «History of the Exponential and Logarithmic Concepts», *Amer. Math. Monthly*, 20, 1913, 5-14, 35-47, 75-84, 107-117, 148-151, 173-182 y 205-210.

—: *A History of the Conceptions of Limits and Fluxions in Great Britain from Newton to Woodhouse*, Open Court, 1919.

—: «The History of Notations of the Calculus», *Annals of Math.*, (2), 25, 1923, 1-46.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teubner, 1898 y 1924, vols. 3 y 4, las secciones correspondientes.
Davis, Philip J.: «Leonhard Euler's Integral: A Historical Profile of the Gamma Function», *Amer. Math. Monthly*, 66, 1959, 849-869.
Euler, Leonhard: *Opera Omnia*, B. G. Teubner y Orell Füssli, 1911-; ver en el capítulo referencias a volúmenes específicos.
Fagnano, Giulio Carlo: *Opere Matematiche*, 3 vols., Albrighi Segati, 1911.
Fuss, Paul H. von: *Correspondance mathématique et physique de quelques célèbres géomètres du XVIII*[ème] *siècle*, 2 vols., 1843, Johnson Reprint Corp., 1967.
Hofmann, Joseph E.: «Über Jakob Bernoullis Beiträge zur Infinitesimalmathematik», *L'Enseignement Mathématique* (2), 2, 61-171, 1956; publicado también separadamente por el Institut de Mathématiques, Ginebra, 1957.
Mittag-Leffler, G.: «An Introducion to the Theory of Elliptic Functions», *Annals of Math.*, (2), 24, 1922-23, 271-351.
Montucla, J. F.: *Histoire des Mathématiques*, A. Blanchard (reimpresión), 1960, vol. 3, pp. 110-380.
Pierpont, James: «Mathematical Rigor, Past and Present», *Amer. Math. Soc Bulletin*, 34, 1928, 23-53.
Struik, D. J.: *A Source Book in Mathematics (1200-1800)*, Harvard University Press, 1969, pp. 333-338, 341-351 y 374-391.

Capítulo 20
SERIES

> Leed a Euler, leed a Euler, él es el maestro de todos nosotros.
>
> P. S. LAPLACE

1. Introducción

Las series fueron consideradas en el siglo XVIII, y lo son hoy todavía, una parte esencial del cálculo infinitesimal. En efecto, Newton consideraba las series como inseparables de su método de fluxiones, ya que la única manera en que podía manejar las funciones algebraicas mínimamente complicadas y las funciones trascendentes era desarrollándolas en serie y derivando o integrando término a término. Leibniz, en sus primeros artículos publicados de 1684 y 1686, también dio importancia a las «ecuaciones generales o indefinidas». Los Bernoulli, Euler y sus contemporáneos confiaban grandemente en el uso de series. Sólo gradualmente, como señalamos en el capítulo precedente, descubrieron los matemáticos cómo trabajar con las funciones elementales en forma cerrada, es decir, como expresiones analíticas simples. No obstante, las series eran la única representación para ciertas funciones y el medio más eficaz para operar con las funciones trascendentes elementales.

Los éxitos obtenidos mediante la utilización de series fueron siendo más numerosos a medida que los matemáticos desarrollaban su disciplina; las dificultades con el nuevo concepto no fueron identi-

ficadas como tales, al menos por un tiempo; las series eran simplemente polinomios infinitos y se podían tratar como tales. Por otro lado, parecía claro, como creían Euler y Lagrange, que toda función podía expresarse en forma de serie.

2. Los primeros trabajos sobre series

Las series, normalmente bajo la forma de progresiones geométricas indefinidas de razón menor que 1, aparecen muy pronto en matemáticas. Aristóteles [1] incluso admitió que tales series poseen una suma. Aparecen esporádicamente entre los matemáticos medievales tardíos, quienes consideraron series para calcular la distancia recorrida por cuerpos móviles cuando la velocidad cambia de un período temporal a otro. Oresme, que había considerado algunas de esas series, incluso demostró en un tratado, *Quæstiones Super Geometrian Euclidis* (h. 1360), que la serie armónica

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots$$

es divergente por el método que se emplea actualmente, a saber, reemplazándola por la serie de términos más pequeños

$$\frac{1}{2}+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)+\cdots$$

y observando que esta última diverge porque podemos obtener tantos grupos de términos, cada uno de ellos de valor 1/2, como queramos. No obstante, no debemos concluir de ello que Oresme o los matemáticos en general comenzaron a distinguir entre series convergentes y divergentes.

En su *Varia Responsa* (1593, *Opera*, 347-435), Vieta dio la fórmula para la suma de una progresión geométrica *infinita*; tomó los *Elementos* de Euclides que la suma de n términos $a_1 + a_2 + \cdots + a_n$ está dada por

$$\frac{s_n - a_n}{s_n - a_1} = \frac{a_1}{a_2}.$$

[1] *Physica*, libro III, cap. 6, 206*b*, 3-33.

Entonces, si $a_1/a_2 > 1$, a_n se aproxima a 0 cuando n se hace infinito, con lo que

$$s_\infty = \frac{a_1^2}{a_1 - a_2}.$$

A mediados del siglo XVII, Gregory de Saint Vincent, en su *Opus Geometricum* (1647), demostró que la paradoja de Aquiles y la tortuga se podía resolver sumando una serie geométrica infinita; la finitud de la suma probaba que Aquiles alcanzaría a la tortuga en un instante y lugar bien definidos. Gregory hizo la primera afirmación explícita de que una serie infinita representa una magnitud, a saber, la suma de la serie, a la que él denominó límite de la serie. Afirma que el «término de una progresión es el final de la serie al que la progresión no alcanza, incluso aunque se continúe hasta el infinito, pero al que se aproxima con un error menor que cualquier intervalo dado». Hizo muchas otras afirmaciones menos exactas y claras, pero realizó contribuciones en la materia e influyó en numerosos discípulos.

Mercator y Newton (cap. 17, sec. 2) descubrieron la serie

$$\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots,$$

de la que se observó que tiene un valor infinito para $x = 2$, mientras que, de acuerdo con el primer miembro, debería valer log 3. Wallis advirtió esta dificultad pero no pudo explicarla. Newton obtuvo muchas otras series para funciones algebraicas y trascendentes. Así, para obtener la serie de arc sen x, en 1666, utilizó el hecho (fig. 20.1) de que el área $OBC = (1/2)$ arc sen x, de donde arc sen $x = \int_0^x \sqrt{1-x^2}\,dx - x\sqrt{1-x^2}/2$. Consiguió el resultado desarrollando en serie el segundo miembro, integrando término a término y combinando las dos series; obtuvo también la serie de arc tg x. En su obra *De Analysi*, de 1669, dio las series de sen x, cos x, *arc sen* x y e^x; algunas de ellas las dedujo de otras invirtiendo una serie, es decir, despejando la variable independiente en términos de la variable dependiente. Su método para lograrlo es tosco e inductivo, pero Newton estaba enormemente satisfecho por haber deducido tantas series.

Collins recibió *De Analysi* en 1669 y comunicó los resultados sobre series a James Gregory el 24 de diciembre de 1670. Este res-

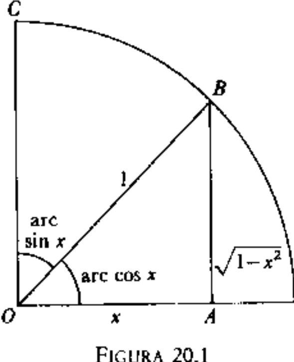


FIGURA 20.1

pondió (Turnbull, *Correspondence*, 1, 52-58 y 61-64) el 15 de febrero de 1671 diciendo que él había obtenido otras series, entre ellas

$$\operatorname{tg} x = x + \frac{x^3}{3} + \frac{2}{15}x^5 + \frac{17}{317}x^7 + \cdots,$$

$$\sec x = 1 + \frac{x^2}{2} + \frac{5}{24}x^4 + \frac{61}{720}x^6 + \cdots.$$

No se sabe cómo obtuvo estas series. También Leibniz obtuvo las series de sen x, cos x y arc tg x, presumiblemente de un modo independiente, en 1673. Las series de las funciones trascendentes constituían el método disponible más fecundo para manejar dichas funciones en las primeras etapas del cálculo infinitesimal y representan una parte importante del trabajo de Newton y Leibniz en este campo.

Estos y otros pensadores que utilizaban el teorema binomial para exponentes negativos y fraccionarios en la obtención de muchas de esas series no sólo pasaban por alto las cuestiones que se plantean en el uso de series, sino que no disponían siquiera de una demostración de dicho teorema; aceptaban también sin ninguna duda que la función que se desarrollaba en serie era efectivamente igual a ésta.

En 1702, Jacques Bernoulli [2] obtuvo las series de sen x y cos x por medio de expresiones que él había obtenido para sen $n\alpha$ en términos de sen α, haciendo tender α a 0 mientras n se hace infinito,

[2] *Opera*, 2, 921-929.

de modo que $n\alpha$ tiende a x mientras que n sen α, que es igual a $n\alpha$ sen α/α, también tiende a x. Wallis había mencionado en la edición latina de su *Algebra* (1693) que Newton había dado de nuevo estas series en 1676; Bernoulli tomó nota de esta observación pero no reconoció la prioridad de Newton; por otra parte, de Moivre dio una demostración de los resultados de Newton en las *Philosophical Transactions* de 1698 [3], y aunque Bernoulli utilizó y se refirió a esta revista en otro trabajo, no dio ninguna indicación de que conociese por esta fuente el trabajo de Newton.

Una de las principales aplicaciones de las series, aparte de su uso en cuestiones de derivación e integración, está en el cálculo de cantidades especiales, tales como π y e, y en las funciones trigonométricas y la logarítmica. Newton, Leibniz, James Gregory, Cotes, Euler y muchos otros estaban interesados en las series con este propósito. Sin embargo, algunas series convergen tan lentamente que en la práctica no son útiles para efectuar cálculos. Así, Leibniz obtuvo en 1674 [4] el célebre resultado

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

Pero serían necesarios alrededor de 100.000 términos para computar π y eso con una precisión como la obtenida por Arquímedes. Análogamente, la serie de $\log(1 + x)$ converge muy lentamente, de modo que sería necesario tomar muchos términos para lograr una precisión de pocos decimales. Esta serie fue transformada de varias maneras a fin de conseguir series que convergiesen más rápidamente; así, James Gregory (*Exercitationes Geometricae*, 1668) obtuvo

$$\frac{1}{2}\log\left(\frac{1+z}{1-z}\right) = z + \frac{1}{3}z^3 + \frac{1}{5}z^5 + \cdots,$$

que se reveló más útil para el cálculo de logaritmos. A lo largo del siglo XVIII fueron muchos los que buscaron transformaciones de una serie en otra que converja más rápidamente; en la sección 4 se da una de esas transformaciones que es debida a Euler.

---

[3] Volumen 20, 190-193.

[4] *Math. Schriften*, 5, 88-92; también *Acta Erud.*, 1682 = *Math. Schriften*, 5, 118-122.

Newton inició otra aplicación más de las series; dada la función implícita $f(x,y) = 0$, para trabajar con la y como función de la x sería deseable disponer de la correspondiente función explícita; puede haber varias de estas funciones explícitas, como resulta evidente del ejemplo trivial de $x^2 + y^2 - 1 = 0$, que tiene las dos soluciones $y = \pm\sqrt{1 - x^2}$, ambas emanando del punto $(1, 0)$. En este caso simple, las dos soluciones se pueden expresar en términos de expresiones analíticas cerradas, pero, en general, cada una de las expresiones de y vendrá dada como una serie en x; no obstante, estas series no son necesariamente series de potencias, en particular si los puntos alrededor de los cuales se busca el desarrollo son puntos singulares ($f_x = f_y = 0$). En su *Method of Fluxions* Newton publicó un esquema para determinar las formas de las distintas series, una para cada solución explícita; su método, que utiliza lo que se conoce como paralelogramo de Newton, muestra cómo determinar los primeros exponentes en una serie de la forma

$$y = a_1 x^m + a_2 x^{m+n} + a_3 x^{m+2n} + \cdots.$$

Los coeficientes de la serie se determinan entonces por el método de coeficientes indeterminados; de hecho, Newton sólo dio ejemplos concretos pero se puede inferir de ellos el método general.

El problema de determinar los exponentes de cada serie es fastidioso, y Taylor, James Stirling y Maclaurin dieron algunas reglas; Maclaurin intentó extenderlas y demostrarlas pero no logró progresos. Una demostración del método de Newton fue llevada a cabo independientemente por Gabriel Cramer y Abraham G. Kästner (1719-1800).

3. Los desarrollos de funciones

Uno de los problemas a los que se enfrentaron los matemáticos de finales del siglo XVII y del XVIII fue el de la interpolación de tablas de valores. Era necesaria una mayor precisión de los valores interpolados de las tablas trigonométricas, logarítmicas y náuticas para ir al paso de los progresos en navegación, astronomía y geografía. El método usual de interpolación (la palabra es de Wallis) se denomina interpolación lineal porque se supone que la función es una función lineal de la variable independiente en el intervalo entre

dos valores conocidos; pero las funciones en cuestión no son lineales y los matemáticos eran conscientes de que se necesitaba un método de interpolación mejor.

El método que vamos a describir fue iniciado por Brigss en su *Arithmetica Logarithmica* (1624), aunque la fórmula clave la dieron James Gregory en una carta a Collins (Turnbull, *Correspondence*, 1, 45-48) del 23 de noviembre de 1670, e, independientemente, Newton. El trabajo de éste aparece en el lema 5 del libro III de los *Principia* y en *Methodus Differentialis*, que, aunque publicado en 1711, fue escrito hacia 1676; el método utiliza lo que se conoce como diferencias finitas y es el primer resultado importante relativo al cálculo de éstas.

Supongamos que $f(x)$ es una función cuyos valores se conoce en $a, a+c, a+2c, a+3c, \cdots, a+nc$. Sean

$$\begin{aligned} \Delta f(a) &= f(a+c) - f(a), \\ \Delta f(a\,c) &= f(a+2c) - f(a+c), \\ \Delta f(a\,2c) &= f(a+3c) - f(a+2c), \\ &\ldots\ldots\ldots \end{aligned}$$

Sean, además

$$\begin{aligned} \Delta^2 f(a) &= \Delta f(a+c) - \Delta f(a) \\ \Delta^3 f(a) &= \Delta^2 f(a+c) - \Delta^2 f(a) \\ &\ldots\ldots\ldots \end{aligned}$$

Entonces la fórmula de Gregory-Newton establece que

$$f(a+h) = f(a) + \frac{h}{c}\Delta f(a) + \frac{\frac{h}{c}\left(\frac{h}{c} - 1\right)}{1 \cdot 2}\Delta^2 f(a) + \cdots. \qquad (1)$$

Newton esbozó una demostración, no así Gregory.

Para calcular $f(x)$ en un valor de x comprendido entre dos valores en los que se conoce $f(x)$ basta hacer simplemente $h = x - a$; el valor calculado no es necesariamente el verdadero valor de la función; lo que da la fórmula es el valor de un polinomio en h que coincide con la verdadera función en los valores especiales $a, a+c, a+2c, \ldots$

La fórmula de Gregory-Newton se aplicó también en integración aproximada. Dada una función $g(x)$ que hay que integrar, quizá para

hallar el área bajo la correspondiente curva, se determinan $g(a)$, $g(a+c)$, $g(a+2c)$, ... y sus diferencias primeras y de orden superior; se sustituyen estos valores en (1) y se obtiene entonces una aproximación polinomial de $g(x)$; y dado que los polinomios, como señala Newton, se integran fácilmente, se consigue así una aproximación a la integral de $g(x)$.

Gregory aplicó también (1) a la función $(1+d)^x$; conocía los valores de esta función en $x=0, 1, 2, 3, \ldots$ y con ellos obtuvo $f(0)=1$, $\Delta f(0)=d$, $\Delta^2 f(0)=d^1$, y así sucesivamente. Haciendo entonces $a=0$, $c=1$ y $h=x-0$ en (1) y utilizando los valores de $f(0)$, $\Delta f(0)$, ..., obtuvo

$$(1+d)^x = 1 + dx + \frac{x(x-1)}{1\cdot 2}d^2 + \frac{x(x-1)(x-2)}{1\cdot 2\cdot 3}d^3 + \cdots. \qquad (2)$$

Así pues, Gregory obtuvo el desarrollo binomial para un x general.

La fórmula de interpolación de Gregory-Newton fue utilizada por Brook Taylor para elaborar el método más potente para desarrollar una función en serie. El teorema binomial, el efectuar la división en una función racional o el método de coeficientes indeterminados son recursos limitados. En su *Methodus Incrementorum Directa et Inversa* (1715), la primera publicación en la que trató del cálculo de diferencias finitas, Taylor derivó el teorema que todavía lleva su nombre y que él mismo había enunciado en 1712. Por cierto, elogia a Newton pero no hace mención del trabajo de Leibniz de 1673 sobre diferencias finitas, a pesar de que Taylor conocía este trabajo.

El teorema de Taylor ya era conocido por James Gregory en 1670 y fue descubierto independientemente algo más tarde por Leibniz, pero ninguno de ellos lo publicó. Jean Bernoulli publicó prácticamente el mismo resultado en las *Acta Eruditorum* de 1694, y aunque Taylor conocía este resultado no se refirió a él. Su propia «demostración» fue de un tipo diferente; lo que hizo equivale a sustituir c por Δx en la fórmula de Gregory-Newton; entonces, por ejemplo, el tercer término del segundo miembro de (1) se convierte en

$$\frac{h(h-\Delta x)}{1\cdot 2}\;\frac{\Delta^2 f(a)}{\Delta x^2}. \qquad (3)$$

Taylor concluía que cuando $\Delta x = 0$, este término se convertía en

$h^2 f''(a)/2!$, de manera que la fórmula completa de Newton-Gregory quedaba como

$$f(a+h)=f(a)+f'(a)h+f''(a)\frac{h^2}{2!}+f'''(a)\frac{h^3}{3!}+\cdots. \qquad (4)$$

Por supuesto, el método de Taylor no era riguroso ni tampoco él se planteó la cuestión de la convergencia.

El teorema de Taylor para $a=0$ se conoce hoy como teorema de Maclaurin. Colin Maclaurin, que sucedió a James Gregory como profesor en Edimburgo, estudió este caso especial en su *Treatise of Fluxions* (1742), afirmando que no era sino un caso especial del resultado de Taylor; históricamente, sin embargo, se le ha atribuido a Maclaurin como un teorema independiente. Por cierto que Sterling dedujo este caso especial para funciones algebraicas en 1717 y para funciones generales en su *Methodus Differentialis* de 1730.

La demostración de Maclaurin de su resultado se basa en el método de coeficientes indeterminados; procede como sigue: sea

$$f(z)=A+Bz+Cz^2+Dz^3+\cdots. \qquad (5)$$

Entonces

$$f'(z)=B+2Cz+3Dz^2+\cdots$$
$$f''(z)=2C+6Dz+\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Haciendo $z=0$ en cada ecuación se determinan los coeficientes A, B, $C, \cdots$. Maclaurin no se preocupó de los problemas de convergencia y procedió a utilizar el resultado.

4. El manejo de las series

Jacques y Jean Bernoulli llevaron a cabo una gran cantidad de trabajos con series. Jacques escribió cinco artículos entre 1689 y 1704 que fueron publicados por su sobrino Nicolaus (1695-1726, hijo de Jean) como suplemento al *Ars Conjectandi* (1713) de Jacques. La mayor parte del trabajo en estos artículos está dedicado al uso de representaciones en serie de funciones con el objeto de derivarlas e

integrarlas y obtener áreas bajo curvas y longitudes de curvas. Aunque estas aplicaciones constituían una importante contribución al cálculo infinitesimal, no eran especialmente originales; sin embargo, merecen destacarse algunos de los métodos que utilizaba para sumar series, pues ilustran la naturaleza del pensamiento matemático del siglo XVIII.

En el primer artículo (1689) [5], Bernoulli comienza con la serie

$$N = \frac{a}{c} + \frac{a}{2c} + \frac{a}{3c} + \cdots, \tag{6}$$

de donde

$$N - \frac{a}{c} = \frac{a}{2c} + \frac{a}{3c} + \frac{a}{4c} + \cdots \tag{7}$$

A continuación, resta (7) de (6) de modo que cada término del segundo miembro de (7) se resta del que tiene encima de (6). Esto da

$$\frac{a}{c} + \frac{a}{1 \cdot 2c} + \frac{a}{2 \cdot 3c} + \frac{a}{3 \cdot 4c} + \cdots, \tag{8}$$

resultado correcto, pero que está incorrectamente deducido porque la serie original es divergente. El mismo afirma que el procedimiento es discutible y que no ha de usarse sin cierta prudencia.

Considera después la serie armónica ordinaria y demuestra que su suma es infinita [6]. Toma los términos

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n^2} \tag{9}$$

y afirma que su suma es mayor que $(n^2 - n) \cdot (1/n^2)$, ya que son $n^2 - n$ términos y cada uno de ellos vale al menos tanto como el último. Pero

$$(n^2 - n) \cdot (1/n^2) = 1 - \frac{1}{n},$$

[5] *Opera*, 1, 375-402.
[6] *Opera*, 1, 392.

con lo que, si añadimos $1/n$ a (9), resulta

$$\frac{1}{n}+\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n^2}>1.$$

Así pues, afirma, podemos formar grupos de términos cada uno de los cuales tiene una suma mayor que 1, y por consiguiente podemos obtener un número finito de términos cuya suma sea tan grande como queramos; en consecuencia, la suma de la serie completa ha de ser infinita. Sucede pues, señala también Bernoulli, que la suma de una serie cuyo «último» término se anula puede ser infinita, contra lo que él opinaba anteriormente y lo que opinaban muchos de los matemáticos del siglo XVIII, incluido Lagrange.

Jean Bernoulli había dado antes una «demostración» diferente de la suma infinita de la serie armónica; se procede así:

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots=\frac{1}{1\cdot 2}+\frac{2}{2\cdot 3}+\frac{3}{3\cdot 4}+\frac{4}{4\cdot 5}+$$

$$=\left(\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\frac{1}{4\cdot 5}+\cdots\right)$$

$$+\left(\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\frac{1}{4\cdot 5}+\cdots\right)$$

$$+\left(\frac{1}{3\cdot 4}+\frac{1}{4\cdot 5}+\cdots\right)=\left(\frac{1}{4\cdot 5}+\cdots\right)+\cdots. \qquad (10)$$

Utilizando ahora (8), con a y c valiendo 1, obtenemos de (10) que

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots=$$

$$1+\left(1-\frac{1}{2}\right)+\left(1-\frac{1}{2}-\frac{1}{6}\right)+\left(1-\frac{1}{2}-\frac{1}{6}-\frac{1}{12}\right)+\cdots$$

$$=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots.$$

Si se pone $A = 1/2 + 1/3 + 1/4 + 1/5 + \cdots$, habríamos probado que $A = 1 + A$, lo que es imposible si A fuese finito.

En los siguientes cuatro estudios sobre series, Jacques Bernoulli hace muchas cosas con tan poca precisión que cuesta creer que reconociese alguna vez la necesidad de tener cautela con la series. Por ejemplo, en el segundo de los artículos (1692) [7] argumenta como sigue: de la fórmula para las progresiones geométricas se tiene que $1 + 1/2 + 1/4 + 1/8 + \cdots = 2$; entonces, tomando 1/3 de ambos miembros se obtiene $1/3 + 1/6 = 1/12 + \cdots = 2/3$; si se toma 1/5 queda $1/5 + 1/10 + 1/20 + \cdots = 2/5$, y así sucesivamente; la suma de los miembros de la izquierda, que es la serie armónica completa, será igual a la suma de los miembros de la derecha, o sea,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = 2 + \frac{2}{3} + \frac{2}{5} + \cdots = 2\left(1 + \frac{1}{3} + \frac{1}{5} + \cdots\right).$$

En consecuencia, la suma de los términos impares es la mitad de la suma de la serie armónica, de donde $1 + 1/3 + 1/5 + \cdots = 1/2 + 1/4 + 1/6 + \cdots$.

En el tercer artículo (1696) [8] escribe

$$\frac{l}{m+n} = \frac{l}{m}\left(1 + \frac{n}{m}\right)^{-1} = \frac{l}{m} - \frac{ln}{m^2} + \frac{ln^2}{m^3} - \cdots;$$

de donde, cuando $n = m$,

$$\frac{l}{2m} = \frac{l}{m} - \frac{l}{m} + \frac{l}{m} + \cdots, \tag{11}$$

lo que describe como una paradoja nada inelegante.

En el segundo artículo sobre series reemplazó el término general por la suma o la diferencia de otros dos términos generales, efectuando después operaciones que le llevaban a determinados resultados; esas sustituciones son legítimas para series absolutamente con-

[7] *Opera*, 1, 517-542.
[8] *Opera*, 2, 745-764.

vergentes, pero no para las condicionalmente convergentes; llegó por ello a resultados erróneos, que él describía también como paradojas.

Uno de los resultados ciertamente interesantes de Jacques se refiere a la serie de recíprocos de las potencias n-ésimas de los números naturales, o sea, la serie $1 + 1/2^n + 1/3^n + 1/4^n + \cdots$; probó que la suma de los términos de lugar impar es a la suma de los términos de lugar par como $2^n - 1$ es a 1; esto es correcto para $n \geqslant 2$, pero no dudó en aplicarlo al caso $n = 1$ y $n = 1/2$, resultado este último que encontró paradójico.

Otro resultado sobre series debido a Jacques Bernoulli afirma que la suma de la serie $1 + 1/\sqrt{2} + 1/\sqrt{3} + \cdots$ es infinita, porque cada término es mayor que el correspondiente de la serie armónica; aquí utilizó con acierto el criterio de comparación.

La serie que provocó la mayor discusión y controversia fue

$$1 - 1 + 1 - 1 + \cdots, \tag{12}$$

que es la (11) cuando l y m son iguales a 1.

Parecía claro que escribiendo la serie en la forma

$$(1 - 1) + (1 - 1) + (1 - 1) + \cdots \tag{13}$$

la suma debería ser 0. Igual de claro resultaría escribiendo la serie como

$$1 - (1 - 1) - (1 - 1) - (1 - 1) \cdots$$

que la suma habría de ser 1. Pero, si se denota la suma de (12) por S, entonces $S = 1 - S$, de donde $S = 1/2$, que es precisamente el resultado de Bernoulli en (11). Guido Grandi (1671-1742), un profesor de matemáticas de la universidad de Pisa, en su pequeño libro *Quadratura Circuli et Hyperbolae* (*La cuadratura de círculos e hipérbolas*, 1703), obtuvo el tercer resultado por otro método; hizo $x = 1$ en el desarrollo

$$\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \cdots \tag{14}$$

y obtuvo

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots.$$

Grandi sostenía por esta razón que 1/2 era la suma de la serie (12); sostenía también que puesto que la suma de (12) en la forma (13) era 0, había logrado probar que el mundo puede ser creado de la nada.

En una carta a Christian Wolf (1678-1754), publicada en las *Acta* [9], Leibniz consideró también la serie (12). Se mostraba de acuerdo con el resultado de Grandi, pero pensaba que era posible obtenerlo sin recurrir a su razonamiento; en su lugar, Leibniz argüía que si toma el primer término, la suma de los dos primeros, la suma de los tres primeros y así sucesivamente, se obtiene 1, 0, 1, 0, 1,· · ·, de modo que 1 y 0 son igualmente probables y por lo tanto habría que tomar su media aritmética, que es también el valor más probable, como valor de la suma. Esta solución fue aceptada por Jacques y Jean Bernoulli, Daniel Bernoulli y, como veremos, por Lagrange; Leibniz concedía que su razonamiento era más metafísico que matemático pero fue más allá, hasta decir que había más de verdad metafísica en la matemática de lo que generalmente se reconocía. Sin embargo, estaba probablemente mucho más influenciado por el razonamiento de Grandi de lo que él mismo se daba cuenta, pues cuando, en correspondencia posterior, Wolf quería concluir que

$$1-2+4-8+16\cdots=\frac{1}{3}$$

$$1-3+9-27+81\cdots=\frac{1}{4}$$

utilizando una extensión del argumento probabilístico de Leibniz, éste le puso objeciones, indicando que las series que tienen suma tienen términos decrecientes, y que (12) es al menos límite de series con términos decrecientes, como resulta evidente de (14) haciendo tender x a 1 por la izquierda.

En realidad, los trabajos sobre series comenzaron en toda su amplitud alrededor de 1730 con Euler, en quien el tema despertó un enorme interés. Había, sin embargo, mucha confusión en su pensamiento; para obtener la suma de

$$1-1+1-1+1-\cdots$$

---

[9] *Acta Erud. Supplementum*, 5, 1713, 264-270 = *Math. Schriften*, 5, 382-387.

Euler decía que, dado que

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots , \tag{15}$$

entonces, cuando $x = 1$,

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots , \tag{16}$$

con lo que aquella suma es 1/2.

Del mismo modo, con $x = -2$ se tiene en (15)

$$\frac{1}{3} = 1 - 2 + 2^2 - 2^3 + \cdots, \tag{17}$$

con lo que la suma de la serie del segundo miembro vale 1/3. Como tercer ejemplo, dado que

$$\frac{1}{(1+x)^2} = (1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + \cdots, \tag{18}$$

se tiene para $x = 1$

$$\frac{1}{4} = 1 - 2 + 3 - 4 + \cdots .$$

Análogamente, como

$$\frac{1-x}{(1+x)^2} = (1-x)(1+x)^{-2} = 1 - 3x + 5x^2 - 7x^3 + \cdots,$$

entonces, para $x = 1$ se tiene

$$0 = 1 - 3 + 5 - 7 + \cdots. \tag{19}$$

Hay multitud de ejemplos de este tipo de razonamientos en sus trabajos.

Haciendo $x = 1$ en (18) se ve que

$$\infty = 1 + 2 + 3 + 4 + 5 + \cdots. \tag{20}$$

Euler aceptaba este resultado; por otra parte, haciendo $x = 2$ en (15) se ve que

$$-1 = 1 + 2 + 4 + 8 + \cdots; \qquad (21)$$

y como el segundo miembro de (21) ha de superar al segundo miembro de (20), resulta que la suma $1 + 2 + 4 + 8 + \cdots$ ha de superar a ∞; pero, de acuerdo con (21), esa suma da -1. Euler concluía que ∞ ha de ser una especie de frontera entre los números positivos y los negativos y que en este aspecto se asemejaba al 0.

A propósito de (19), Nicolaus Bernoulli (1687-1759) afirmó, en una carta a Euler de 1743, que la suma de esta serie $1 - 3 + 5 - 7 + \cdots$ es $-\infty(-1)^{\infty}$; decía que el resultado de Euler (0) era una contradicción insoluble. También observó Bernoulli que haciendo $x = 2$ en (15) se tiene

$$-1 = 1 + 2 + 4 + 8 + \cdots;$$

y que de

$$\frac{1}{1 - x - x^2} = 1 + x + 2x^2 + 3x^3 + \cdots, \qquad (22)$$

para $x = 1$ se tiene

$$-1 = 1 + 1 + 2 + 3 + \cdots.$$

El hecho de que dos series diferentes diesen -1 era también una contradicción insoluble, pues en otro caso se podrían igualar ambas series.

Euler indicó ciertamente en un artículo que las series sólo pueden utilizarse para aquellos valores de x para los que convergen, pero exactamente en el mismo artículo [10] concluía que

$$\cdots + \frac{1}{x^2} + \frac{1}{x} + 1 + x + x^2 + x^3 + \cdots = 0. \qquad (23)$$

Su razonamiento era que

$$\frac{x}{1 - x} = x + x^2 + x^3 + \cdots$$

[10] *Comm. Acad. Sci. Petrop.*, 11, 1739, 116-127, pub. 1750 = *Opera*, (1), 14, 350-363.

y

$$\frac{x}{x-1} = \frac{1}{1-\frac{1}{x}} = 1 + \frac{1}{x} + \frac{1}{x^2} + \frac{1}{x^3} + \cdots$$

Pero la suma de los dos primeros miembros es 0, mientras que la suma de los segundos es la serie original.

En un artículo anterior [11], Euler comenzaba con la serie

$$y = \text{sen}\, x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \qquad (24)$$

o bien

$$1 - \frac{x}{y} + \frac{x^3}{3!y} - \frac{x^5}{5!y} + \cdots = 0. \qquad (25)$$

Por medio de consideraciones algebraicas aplicadas a (25) en tanto que *polinomio de grado infinito*, y utilizando el teorema sobre la relación entre raíces y coeficientes de una ecuación algebraica, Euler probó que [12]:

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}$$

$$\frac{1}{1^3} - \frac{1}{3^3} - \frac{1}{5^3} - \cdots = \frac{\pi^3}{32}$$

$$\frac{1}{1^4} + \frac{1}{3^4} + \frac{1}{5^4} + \cdots = \frac{\pi^4}{96}$$

$$\frac{1}{1^5} - \frac{1}{3^5} + \frac{1}{5^5} - \cdots = \frac{5\pi^5}{1536}$$

$$\frac{1}{1^6} + \frac{1}{3^6} + \frac{1}{5^6} + \cdots = \frac{\pi^6}{960}$$

....

[11] *Comm. Acad. Sci. Petrop.*, 7, 1734/35, 123-134, pub. 1740 = *Opera*, (1), 14, 73-86.

[12] Euler utilizó el símbolo p en lugar de π hasta 1739; π había sido introducido por William Jones en 1706.

En el mismo artículo dio por primera vez el producto infinito

$$\operatorname{sen} s = \left(1 - \frac{s^2}{\pi^2}\right)\left(1 - \frac{s^2}{4\pi^2}\right)\left(1 - \frac{s^2}{9\pi^2}\right)\cdots. \qquad (26)$$

Su argumento fue simplemente que sen s tiene por ceros $\pm\pi$, $\pm 2\pi, \cdots$ (descarta la raíz 0), y en consecuencia, lo mismo que cualquier polinomio, ha de tener un factor lineal correspondiente a cada un de sus raíces. (En 1743 [13] y en su *Introductio* [14], dio otra deducción para hacer frente a las críticas.) Trató el segundo miembro de (26) como un polinomio, que igualó a cero, y utilizando de nuevo la relación entre las raíces y los coeficientes dedujo que

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{90}$$

$$\frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \cdots = \frac{\pi^4}{6},$$

....

y sumas análogas para potencias pares superiores del denominador.

En un artículo posterior [15], Euler obtuvo uno de sus triunfos más brillantes,

$$\sum_{v=1}^{\infty} \frac{1}{v^{2n}} = (-1)^{n-1} \frac{(2\pi)^{2n}}{2(2n)!} B_{2n},$$

donde los B_{2n} son los números de Bernoulli (ver más abajo). En realidad, la relación con los números de Bernoulli la estableció Euler un poco más tarde en su *Institutiones* de 1755 [16]. También calculó

---

[13] *Opera*, (1), 14, 138-155.
[14] *Opera*, (1), 8, 168.
[15] *Comm. Acad. Sci. Petrop.*, 12, 1740, 53-96, pub. 1750 = *Opera*, (1), 14, 407-462.
[16] Parte II, cap. 5, ¶124 = *Opera*, (1), 10, 327.

en el artículo de 1740 la suma $\Sigma_{v=1}^{\infty}(1/v^n)$ para los primeros valores impares de n, aunque no obtuvo la expresión general para todo n impar.

Euler también trabajó sobre series armónicas, es decir, series tales que los recíprocos de sus términos están en progresión aritmética. En particular, demostró [17] cómo se puede sumar un número finito de términos de la serie armónica ordinaria utilizando la función logarítmica. Comienza con

$$\log\left(1+\frac{1}{x}\right)=\frac{1}{x}-\frac{1}{2x^2}+\frac{1}{3x^3}-\frac{1}{4x^4}+\cdots, \tag{27}$$

de donde

$$\frac{1}{x}=\log\left(\frac{x+1}{x}\right)+\frac{1}{2x^2}-\frac{1}{3x^3}+\frac{1}{4x^4}-\cdots.$$

Tomando ahora $x=1, 2, 3, \ldots, n$, se tiene

$$\frac{1}{1}=\log 2+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\cdots$$

$$\frac{1}{2}=\log\frac{3}{2}+\frac{1}{2\cdot 4}-\frac{1}{3\cdot 8}+\frac{1}{4\cdot 16}-\frac{1}{5\cdot 32}+\cdots$$

$$\frac{1}{3}=\log\frac{4}{3}+\frac{1}{2\cdot 9}-\frac{1}{3\cdot 27}+\frac{1}{4\cdot 81}-\frac{1}{5\cdot 243}+\cdots$$

....

$$\frac{1}{n}=\log\frac{n+1}{n}+\frac{1}{2n^2}-\frac{1}{3n^3}+\frac{1}{4n^4}-\frac{1}{5n^5}+\cdots.$$

Sumando, y observando que cada término logarítmico es la diferencia de dos logaritmos, se llega a

[17] *Comm. Acad. Sci. Petrop.*, 7, 1734/35, 150-161, pub. 1740 = *Opera*, (1), 14, 87-100.

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}=\log(n+1)$$

$$+\frac{1}{2}\left(1+\frac{1}{4}+\frac{1}{9}+\cdots+\frac{1}{n^2}\right)-\frac{1}{3}\left(1+\frac{1}{8}+\frac{1}{27}+\cdots+\frac{1}{n^3}\right)$$

$$+\frac{1}{4}\left(1+\frac{1}{16}+\frac{1}{81}+\cdots+\frac{1}{n^4}\right)-\cdots,$$

o bien

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}=\log(n+1)+C, \qquad (28)$$

donde C representa la suma del conjunto infinito de sumas aritméticas finitas. El valor de C fue calculado aproximadamente por Euler (depende de n, pero para n grande el resultado no se ve muy afectado por el valor de n), obteniendo 0,577218. Esta C se conoce actualmente como constante de Euler y se denota por γ; hoy en día se obtiene una representación más precisa de γ de la manera que sigue: restando $\log n$ de ambos miembros de (28) y teniendo en cuenta que $\log(n+1)-\log n=\log(1+1/n)$ tiende a 0 cuando $n\to\infty$ se deduce que

$$\gamma=\lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n\right). \qquad (29)$$

No se ha encontrado, por cierto, una forma más simple que (29) para la constante de Euler, mientras que disponemos de varias expresiones para π y e; no se sabe, por otra parte, si γ es racional o irracional.

En su artículo «De Seriebus Divergentibus» [18], Euler investigó la serie divergente

$$y=x-(1!)x^2+(2!)x^3-(3!)x^4+\cdots. \qquad (30)$$

[18] *Novi Acad. Sci. Petrop.*, 5, 1754/5, 205-237, pub. 1760 = *Opera*, (1), 14, 585-617.

Formalmente, esta serie satisface la ecuación diferencial

$$x^2y' + y = x. \tag{31}$$

Pero esta ecuación diferencial tiene el factor integrante $x^2e^{-1/x}$, de modo que

$$y = e^{1/x}\int_0^x \frac{e^{-1/t}}{t}\,dt \tag{32}$$

es una solución que se puede demostrar por la regla de L'Hospital que se anula con x. Euler consideró la serie (30) como el desarrollo en serie de la función (32) y (32) como la suma de la serie (30). De hecho, hace $x = 1$ y obtiene

$$1 - 1 + 2! - 3! + 4! - \cdots = e\int_0^1 \frac{e^{-1/t}}{t}\,dt.$$

El hecho notable acerca de la serie (30) es que puede utilizarse para obtener buenas aproximaciones numéricas de la función (32) gracias a que, dado un valor de x, si despreciamos todos los términos de la serie a partir de uno determinado, se puede demostrar que el valor absoluto del resto es menor que el valor absoluto del primer término que se desprecia; Euler sacaba partido así de las series divergentes. El pleno significado de lo que permiten realizar estas series divergentes no fue apreciado hasta 150 años más tarde. (Ver. cap. 47.)

Hay que hacer notar otro resultado célebre de Euler en el campo de las series. En su *Ars Conjectandi*, Jacques Bernoulli, tratando el tema de la probabilidad, introdujo los hoy ampliamente utilizados números de Bernoulli. Buscaba una fórmula para las sumas de las potencias enteras positivas de los números enteros y dio sin demostración la siguiente fórmula:

$$\sum_{k=1}^{n} k^c = \frac{1}{c+1}n^{c+1} + \frac{1}{2}n^c + \frac{c}{2}B_2n^{c-1} + \frac{c(c-1)(c-2)}{2\cdot 3\cdot 4}B_4n^{c-3}$$

$$+\frac{c(c-1)(c-2)(c-3)(c-4)}{2\cdot 3\cdot 4\cdot 5\cdot 6}B_6n^{c-5} + \cdots. \tag{33}$$

La suma termina en la última potencia positiva de n. Los B_2, B_4, $B_6, \cdots$ son los números de Bernoulli

$$B_2 = \frac{1}{6},\ B_4 = -\ \frac{1}{30},\ B_6 = \frac{1}{42},\ B_8 = -\ \frac{1}{30},\ B_{10} = \frac{5}{66}, \cdots \qquad (34)$$

Bernoulli dio también la relación de recurrencia que permite calcular esos coeficientes.

El resultado de Euler, la fórmula de sumación de Euler-Maclaurin, es una generalización [19]. Sea $f(x)$ una función real de la variable real x. Entonces (en notación moderna) la fórmula reza así

$$\sum_{i=0}^{n} f(i) = \int_0^n f(x)\ dx - \frac{1}{2}[f(n) - f(0)]$$

$$+ \frac{B_2}{2!}[f'(n) - f'(0)] + \frac{B_4}{4!}[f'''\ (n) - f'''(0)]\ + \cdots$$

$$+ \frac{B_{2k}}{(2k)!}\ [f^{(2k-1)}\ (n) - f^{(2k-1)}\ (0)] + R_k \qquad (35)$$

donde

$$R_k = \int_0^n\ f^{(2k+1)}(x)\ P_{2k+1}(x)\ dx. \qquad (36)$$

Aquí n y k son enteros positivos; $P_{2k+1}(x)$ es el polinomio $(2k+1)$-ésimo de Bernoulli (que aparece también en su obra *Ars Conjectandi*) y está dado por

$$P_k(x) = \frac{x^k}{k!} + \frac{B_1}{1!}\ \frac{x^{k-1}}{(k-1)!} + \frac{B_2}{2!}\ \frac{x^{k-2}}{(k-2)!} + \cdots + \frac{B_k}{k!}, \qquad (37)$$

---

[19] *Comm. Acad. Sci. Petrop.*, 6, 1732/3, 68-97, pub. 1738 = *Opera*, (1), 14, 42-72, y *Comm. Acad. Sci. Petrop.*, 8, 1736, 147-158, pub. 1741 = *Opera*, (1), 14, 124-137.

donde $B_1 = -1/2$ y $B_{2k+1} = 0$ para $k = 1, 2, \cdots$. La serie

$$\sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} [f^{(2k-1)}(n) - f^{(2k-1)}(0)] \tag{38}$$

es divergente para casi todas las $f(x)$ que se presentan en las aplicaciones. Sin embargo, el resto R_k es menor que el primer término que se desprecia, con lo que la serie de (35) proporciona una útil aproximación de

$$\sum_{i=0}^{n} f(i).$$

Los números de Bernoulli B_i se definen hoy, con frecuencia, mediante una relación dada más tarde por Euler [20], a saber,

$$t(e^t - 1)^{-1} = \sum_{i=0}^{\infty} B_i \frac{t^i}{i!}. \tag{39}$$

Independientemente de Euler, Maclaurin [21] llegó a la misma fórmula de sumación (35), pero por un método algo más fundamentado y más próximo al que utilizamos en la actualidad. El resto fue calculado y tratado seriamente por primera vez por Poisson [22].

Euler también introdujo [23] una transformación de series que todavía se conoce y utiliza. Dada una serie $\Sigma_{n=0}^{\infty} b_n$, la escribió en la forma $\Sigma_{n=0}^{\infty} (-1)^n a_n$ y mediante un cierto número de pasos algebraicos formales probó que

$$\sum_{n=0}^{\infty} (-1)^n a_n = \sum_{n=0}^{\infty} (-1)^n \frac{\Delta^n a_0}{2^{n+1}}, \tag{40}$$

---

[20] *Opera*, (10), 14, 407-462.
[21] *Treatise of Fluxions*, 1742, p. 672.
[22] *Mém. de l'Acad. des Sci., Inst. France*, 6, 1823, 571-602, pub. 1827.
[23] *Inst. Cal. Diff.*, 1755, p. 281.

donde Δ^n denota la diferencia finita n-ésima (sec. 3). La ventaja de esta transformación, en términos modernos, consiste en convertir una serie convergente en otra que sea más rápidamente convergente. Para Euler, sin embargo, que normalmente no distinguía entre series convergentes y divergentes, la transformación podía también convertir series divergentes en convergentes. Si se aplica (40) a

$$1 - 1 + 1 - 1 + \cdots, \tag{41}$$

entonces el segundo miembro de (40) da 1/2. Análogamente, para la serie

$$1 - 2 + 2^2 - 2^3 + 2^4 - \cdots \tag{42}$$

(40) da

$$\sum_{n=0}^{\infty} (-1)^n 2^n = \frac{1}{2}(1) + \frac{1}{4}(-1) + \frac{1}{8}(1) - \frac{1}{16}(-1) \cdots = \frac{1}{3}. \tag{43}$$

Estos resultados son, naturalmente, los mismos que Euler obtuvo antes (ver (16) y (17)) tomando para la suma de la serie el valor de la función de la que se deriva la serie.

Ha de quedar claro el espíritu de los métodos de Euler: él es el gran manipulador que señaló el camino para miles de resultados establecidos más tarde rigurosamente.

Hay que mencionar otra serie célebre. En su *Methodus Differentialis* [24], James Stirling consideró la serie que hoy escribimos en la forma

$$\log n! = \left(n + \frac{1}{2}\right) \log n - n + \log \sqrt{2\pi} + \frac{B_2}{1 \cdot 2}\, \frac{1}{n} + \frac{B_4}{3 \cdot 4}\, \frac{1}{n^9} + \cdots + \frac{B_{2k}}{(2k-1)(2k)}\, \frac{1}{n^{2k-1}} + \cdots, \tag{44}$$

[24] 1730, p. 135.

lo que es equivalente a

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \exp\left[\frac{B_2}{1\cdot 2}\frac{1}{n} + \cdots + \frac{B_{2k}}{(2k-1)(2k)}\frac{1}{n^{2k-1}} + \cdots\right]. \quad (45)$$

Stirling dio los cinco primeros coeficientes y una fórmula de recurrencia para determinar los siguientes. Aunque la serie para log $n!$ es divergente, Stirling calculó $\log_{10}$ (1000!), que es 2567 más un decimal, con diez decimales exactos utilizando sólo unos pocos términos de su serie. De Moivre dio en 1730 *(Miscellanea Analytica)* una fórmula similar. Para n grande se tiene $n! \sim (n/e)^n \sqrt{2\pi n}$, resultado que, aunque debido a De Moivre, se conoce como aproximación de Stirling.

5. Series trigonométricas

Los matemáticos del siglo XVIII también trabajaron mucho con las series trigonométricas, especialmente en su teoría astronómica. La utilidad de tales series en astronomía es evidente si se tiene en cuenta que son funciones periódicas y que los fenómenos astronómicos son en buena medida periódicos. Ese trabajo representó el origen de un vasto campo cuyo pleno significado no fue apreciado en el siglo XVIII. El problema con el que se inició el uso de series trigonométricas fue el de la interpolación, en particular, la determinación de las posiciones de los planetas intermedias a las obtenidas por observación. Las mismas series fueron también introducidas en los primeros trabajos sobre ecuaciones en derivadas parciales (ver cap. 22), pero curiosamente las dos líneas de pensamiento se mantuvieron separadas a pesar de que trabajaron en ambos problemas las mismas personas.

Por serie trigonométrica se entiende una serie de la forma

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx), \quad (46)$$

donde a_n y b_n son constantes. Si tal serie representa una función $f(x)$,

entonces

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \operatorname{sen} nx \, dx \tag{47}$$

para $n = 0, 1, 2, \ldots$. La consecución de estas fórmulas para los coeficientes fue uno de los resultados capitales de la teoría, aunque no diremos nada por el momento acerca de las condiciones bajo las cuales esos son necesariamente los valores de a_n y b_n.

Euler había abordado en una fecha tan temprana como 1729 el problema de la interpolación, o sea, dada una función $f(x)$ cuyos valores para $x = n$, n entero positivo, están prescritos, hallar $f(x)$ para otros valores de x. En 1747 aplicó el método que había desarrollado a una función que aparece en la teoría de perturbaciones planetarias y obtuvo una representación en serie trigonométrica de la función; publicó dicho método en 1753 [25].

En primer lugar, Euler atacó el problema cuando las condiciones dadas son $f(n) = 1$ para todo n y buscó una función periódica que valiese 1 para x entero. Su razonamiento es interesante porque ilustra el análisis de la época. Pone $f(x) = y$ y, por el teorema de Taylor, escribe

$$f(x+1) = y + y' + \frac{1}{2}y'' + \frac{1}{6}y''' + \cdots \tag{48}$$

Como $f(x+1)$ es igual a $f(x)$, y satisfará la ecuación diferencial lineal de orden infinito

$$y' + \frac{1}{2}y'' + \frac{1}{6}y''' + \cdots = 0. \tag{49}$$

Aplica ahora el método para resolver ecuaciones diferenciales ordinarias lineales publicado en 1743 (ver cap. 21); es decir, plantea la ecuación auxiliar

$$z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \cdots = 0, \tag{50}$$

[25] *Novi Comm. Acad. Sci. Petrop.*, 3, 1750/51, 36-85, pub. 1753 = *Opera*, (1), 14, 463-515.

ecuación que, a la vista de la serie de e, es

$$e^z - 1 = 0.$$

A continuación, determina las raíces de esta última ecuación; para ello, comienza con la ecuación polinómica de grado n

$$\left(1 + \frac{z}{n}\right)^n - 1 = 0;$$

de acuerdo con un teorema demostrado independientemente por Cotes (1722) y Euler en su *Introductio* [26], el polinomio del primer miembro tiene el factor lineal z y los factores cuadráticos

$$\left(1 + \frac{z}{n}\right)^2 - 2\left(1 + \frac{z}{n}\right)\cos\frac{2k\pi}{n} + 1, \qquad k = 1, 2, \cdots, < \frac{n}{2}.$$

En virtud de la identidad trigonométrica para sen z en términos de cos 2z, estos factores son los mismos que

$$4\left(1 + \frac{z}{n}\right)\operatorname{sen}^2\frac{k\pi}{n} + \frac{z^2}{n^2}.$$

Las raíces de (50) no quedan afectadas si dividimos cada factor por 4 $\operatorname{sen}^2 k\pi/n$ (para el correspondiente k), con lo que los factores cuadráticos serán

$$1 + \frac{z}{n} + \frac{z^2}{4n^2\operatorname{sen}^2\dfrac{k\pi}{n}}.$$

Para $n = \infty$, el término z/n es 0; reemplazando la cantidad sen $k\pi/n$ por $k\pi/n$, los factores se convierten en

$$1 + \frac{z^2}{4k^2\pi^2}.$$

A cada uno de esos factores en la ecuación auxiliar (50) le corresponden las raíces $z = \pm i2k\pi$, y, en consecuencia, la solución

$$\alpha_k \operatorname{sen} 2k\pi x + A_k \cos 2k\pi x$$

[26] Volumen 1, cap. 14.

de (49). El factor lineal z mencionado antes da lugar a una solución constante y, dado que $f(0) = 1$ es una condición inicial, Euler obtiene finalmente

$$y = 1 + \sum_{k=1}^{\infty} \{\alpha_k \text{ sen } 2k\pi x + A_k(\cos 2k\pi x - 1)\},$$

donde los coeficientes α_k y A_k han de ser tales que se satifaga la condición de que $f(n) = 1$ para todo n.

Este artículo contiene también un resultado que es formalmente idéntico a lo que llegaría a conocerse como desarrollo de Fourier de una función arbitraria, así como la determinación de los coeficientes mediante integrales. Concretamente, Euler demostró que la solución general de la ecuación funcional

$$f(x) = f(x - 1) + X(x)$$

es

$$f(x) = \int_0^x X(\xi)\, d\xi + 2 \sum_{n<1}^{\infty} \cos 2n\pi x \int_0^x X(\xi) \cos 2n\pi\xi\, d\xi$$

$$+ 2 \sum_{n=1}^{\infty} \text{sen } 2n\pi x \int_0^x X(\xi) \text{ sen } 2n\pi\xi\, d\xi.$$

Tenemos aquí una función expresada en serie trigonométrica en el año 1750-51. Euler sostenía que la suya era la solución más general del problema de interpolación, con lo que, de ser así, incluiría ciertamente la representación de polinomios con series trigonométricas; sin embargo, Euler rechazó esto en sus argumentos sobre el problema de la cuerda vibrante y otros relacionados, según veremos en el capítulo 22.

En 1754, D'Alembert [27] consideró el problema de desarrollar el

[27] *Recherches sur différens points importans du système du monde*, 1754, vol. II, p. 66.

recíproco de la distancia entre los planetas en una serie de cosenos de los múltiplos del ángulo que forman los rayos que van del origen a los planetas, y aquí también podemos encontrar las expresiones en integrales definidas para los coeficientes de las series de Fourier.

En otro trabajo, Euler obtiene representaciones de funciones en series trigonométricas de una manera completamente diferente [28]. Comienza con la serie geométrica

$$\sum_{n=0}^{\infty} a^n (\cos x + i \operatorname{sen} x)^n, \quad i = \sqrt{-1}$$

y sumándola obtiene

$$\frac{1}{1 - a\,(\cos x + i \operatorname{sen} x)}.$$

Utiliza entonces fórmulas bien conocidas para reemplazar las potencias de cos x y sen x por cos nx y sen nx (lo que equivale al teorema de De Moivre) y obtiene

$$\frac{1}{1 - a\,(\cos x + i \operatorname{sen} x)} = \sum_{n=0}^{\infty} a^n (\cos nx + i \operatorname{sen} nx).$$

Multiplicando numerador y denominador del primer miembro por el conjugado del denominador, pasando el término correspondiente a $n = 0$ al primer miembro, dividiendo todo por a y separando las partes real e imaginaria, obtiene

$$\frac{a \cos x - a^2}{1 - 2a \cos x + a^2} = \sum_{n=1}^{\infty} a^n \cos nx$$

$$\frac{a \operatorname{sen} x}{1 - 2a \cos x + a^2} = \sum_{n=1}^{\infty} a^n \operatorname{sen} nx \cdot$$

[28] *Novi Comm. Acad. Sci. Petrop.*, 5, 1754/5, 164-204, pub. 1760 = *Opera*, (1), 14, 542-584; ver también *Opera*, (1), 15, 435-497, para otro método.

Hasta aquí, sus resultados no son sorprendentes. Euler hace ahora $a = \pm 1$ y obtiene, por ejemplo,

$$\frac{1}{2} = 1 \pm \cos x + \cos 2x \pm \cos 3x + \cos 4x \pm \cdots. \qquad (51)$$

(La serie es en realidad divergente.) A continuación integra y obtiene

$$\frac{\pi - x}{2} = \operatorname{sen} x + \frac{1}{2} \operatorname{sen} 2x + \frac{1}{3} \operatorname{sen} 3x + \cdots, \qquad (52)$$

(que se verifica para $0 < x < \pi$ y vale 0 para $x = 0$ y $x = \pi$) y

$$\frac{x}{2} + \operatorname{sen} x - \frac{1}{2} \operatorname{sen} 2x + \frac{1}{3} \operatorname{sen} 3x - \frac{1}{4} \operatorname{sen} 4x + \cdots, \qquad (53)$$

(que converge en $-\pi < x < \pi$). Integrando esta última expresión, con evaluación en $x = 0$ para determinar la constante de integración, se obtiene

$$\frac{x^2}{4} - \frac{\pi^2}{4} = -\cos x + \frac{1}{4} \cos 2x - \frac{1}{9} \cos 3x - \frac{1}{16} \cos 4x + \cdots. \qquad (54)$$

Euler pensaba que estas dos últimas series [que son convergentes en $(-\pi < x < \pi)$] representaban las respectivas funciones para todo valor de x. Además, derivando sucesivamente (51), Euler deducía que

$$\operatorname{sen} x \pm 2 \operatorname{sen} 2x + 3 \operatorname{sen} 3x \pm \cdots = 0$$

$$\cos x \pm 4 \cos 2x + 9 \cos 3x \pm \cdots = 0$$

y otras ecuaciones del mismo tipo. Daniel Bernoulli, que había obtenido también desarrollos tales como (52), (53) y (54), admitió que las series sólo representaban las funciones para cierto rango de valores de la x.

En 1757, cuando estudiaba las perturbaciones causadas por el sol, Clairaut[29] dio un paso mucho más audaz. Afirma que va a repre-

[29] *Hist. de l'Acad. des Sci., Paris,* 1754, 545 y sigs., pub. 1759.

sentar *cualquier* función en la forma

$$f(x) = A_0 + 2\sum_{n=1}^{\infty} A_n \cos nx. \tag{55}$$

Considera la cuestión como un problema de interpolación y por ello utiliza los valores de la función en

$$\frac{2\pi}{k}, \ \frac{4\pi}{k}, \ \frac{6\pi}{k}, \ldots$$

obteniendo después de ciertas manipulaciones

$$A_0 = \frac{1}{k}\sum_{\mu} f\left(\frac{2\mu\pi}{k}\right)$$

$$A_n = \frac{1}{k}\sum_{\mu} f\left(\frac{2\mu\pi}{k}\right) \cos \frac{2\mu n\pi}{k}.$$

Haciendo tender k a infinito, Clairaut llega a

$$A_n = \frac{1}{2\pi}\int_0^{2\pi} f(x) \cos nx \, dx, \tag{56}$$

que es la fórmula correcta para A_n.

Lagrange, en su investigación sobre la propagación del sonido [30], obtuvo la serie (51) y defendió el hecho de que su suma es 1/2. Con todo, ni Euler ni Lagrange hicieron observación alguna sobre el hecho notable de que habían expresado funciones no periódicas en forma de serie trigonométrica, aunque algo más tarde sí lo advirtieron en otro contexto. D'Alembert había dado a menudo el ejemplo de $x^{2/3}$ como función que no podía desarrollarse en serie trigonométrica, pero Lagrange le demostró en una carta [31] de 15 de agosto de 1768 que $x^{2/3}$ sí podía desarrollarse en la forma

$$x^{2/3} = a + b \cos 2x + c \cos 4x + \cdots.$$

[30] *Misc. Taur.*, 1, 1759 = *Œuvres*, 1, 110.

[31] *Lagrange, Œuvres*, 13, 116.

D'Alembert le objetó y proporcionó argumentos en contra, tales como que las derivadas de ambos miembros no eran iguales para $x = 0$. Asimismo, por el método de Lagrange se podría expresar sen x en serie de cosenos, a pesar de que sen x es una función impar y el segundo miembro sería una función par. La cuestión no sería resuelta en el siglo XVIII.

En 1777[32], Euler, trabajando en un problema de astronomía, obtuvo de hecho los coeficientes de una serie trigonométrica utilizando la ortogonalidad de las funciones trigonométricas, el método utilizado en la actualidad. O sea, de

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos \frac{k\pi x}{l} \tag{57}$$

dedujo que

$$a_k = \frac{2}{l}\int_0^l f(s) \cos \frac{k\pi s}{l}\, ds.$$

Euler había obtenido esto en el artículo inmediatamente anterior de una manera algo complicada, pero después se dio cuenta de que podía obtenerlo directamente multiplicando los dos miembros de (57) por cos $(\nu\pi x/l)$, integrando término a término y aplicando las relaciones

$$\int_0^l \cos \frac{\nu\pi x}{l} \cos \frac{k\pi x}{l}\, dx = \begin{cases} 0 & \text{si } \nu \neq k \\ l/2 & \text{si } \nu = k \neq 0 \\ l & \text{si } \nu = k = 0. \end{cases}$$

Todos estos trabajos sobre series trigonométricas estaban impregnados por la paradoja de que, si bien se representaba en serie trigonométrica toda suerte de funciones, Euler, D'Alembert y Lagrange nunca abandonaron la opinión de que ello no era posible para funciones *arbitrarias*. La paradoja se explica parcialmente por el hecho de que las series trigonométricas se suponía que eran válidas cuando otro indicio, en ciertos casos de naturaleza física, parecía asegurarlo;

---

[32] *Nova Acta Acad. Sci. Petrop.*, 11, 1793, 114-132, pub. 1798 = *Opera*, (1), 16, parte 1, 333-355.

en ese caso, ellos se consideraban justificados para dar por válida la serie y deducir las fórmulas para los coeficientes. Esta cuestión de si cualquier función se puede representar por una serie trigonométrica se convertiría en una cuestión clave.

6. Fracciones continuas

Ya hemos indicado (cap. 13, sec. 2) el uso de fracciones continuas para obtener aproximaciones de números irracionales. Euler abordó este tema y, en el primer artículo sobre él [33], titulado «De Fractionibus Continuis», dedujo un buen número de resultados interesantes, tales como que todo número racional se puede expresar como una fracción continua finita. Dio después los desarrollos

$$e-1=1+\frac{1}{1+}\ \frac{1}{2+}\ \frac{1}{1+}\ \frac{1}{1+}\ \frac{1}{4+}\ \frac{1}{1+}\ \frac{1}{1+}\ \frac{1}{6+}\cdots,$$

el cual ya había aparecido en un artículo de Cotes en las *Philosophical Transactions* de 1714, y

$$\frac{e+1}{e-1}=2+\frac{1}{6+}\ \frac{1}{10+}\ \frac{1}{14+}\cdots$$

Euler probó, en esencia, que e y e^2 son irracionales.

En su *Introductio* (cap. 18), Euler estableció los fundamentos de una teoría de fracciones continuas, mostrando allí como pasar de una serie a una fracción continua que la representa, y recíprocamente.

El trabajo de Euler sobre fracciones continuas fue utilizado por Johann Heinrich Lambert (1728-77), colega de Euler y Lagrange en la Academia de Ciencias de Berlín, para demostrar [34] que si x es un número racional distinto de cero, entonces e^x y $\operatorname{tg} x$ no pueden ser racionales; demostró de ese modo no sólo que e^x, para x entero positivo, es irracional, sino que todos los números racionales tienen logaritmo natural (base e) irracional. Del resultado para $\operatorname{tg} x$ se sigue que, dado que $\operatorname{tg}(\pi/4)=1$, ni $\pi/4$ ni π pueden ser racionales. Lam-

---

[33] *Comm. Acad. Sci. Petrop.*, 9, 1737, 98-137, pub. 1744 = *Opera*, (1), 14, 187-215.

[34] *Hist. de l'Acad. de Berlin*, 1761, 265-322, pub. 1768 = *Opera*, 2, 112-159.

bert demostró, de hecho, la convergencia del desarrollo en fracción continua para tg x.

Lagrange[35] utilizó fracciones continuas para hallar aproximaciones de las raíces irracionales de ecuaciones, y, en otro artículo de la misma revista[36], obtuvo soluciones aproximadas de ecuaciones diferenciales en forma de fracciones continuas. En el artículo de 1768, Lagrange demostró el recíproco de un teorema que Euler había establecido en su artículo de 1744; dicho recíproco afirma que una raíz real de una ecuación cuadrática es una fracción continua periódica.

7. El problema de la convergencia y la divergencia

Hoy sabemos que los trabajos del siglo XVIII sobre series eran en buena parte formales y que la cuestión de la convergencia y la divergencia no se tomaba ciertamente muy en serio, aunque tampoco fue totalmente pasada por alto.

Newton[37], Leibniz, Euler e incluso Lagrange consideraban las series como una extensión del álgebra de polinomios y así difícilmente advertían que estaban introduciendo nuevos problemas al extender las sumas a un número infinito de términos. No estaban, pues, plenamente equipados para hacer frente a los problemas que las series infinitas les plantearon; con todo las dificultades evidentes que iban surgiendo les indujeron a plantearse estas cuestiones. Lo que llama la atención es que la solución correcta de las paradojas y otras dificultades fue expuesta con frecuencia, pero con igual frecuencia pasada por alto.

La distinción entre convergencia y divergencia había sido ya tenida en cuenta por algunos pensadores del siglo XVII. En 1668, lord Brouncker, tratando de la relación entre log x y el área bajo $y = 1/x$, demostró la convergencia de las series de log 2 y log 5/4 por comparación con una serie geométrica. Newton y James Gregory, quienes hicieron mucho uso de valores numéricos de series para calcular tablas de logaritmos y de otras funciones y para evaluar integrales, eran conscientes de que las sumas de las series podían ser finitas o

[35] *Nouv. Mém. de l'Acad. de Berlin*, 23, 1767, 311-352, pub. 1769 = *Œuvres*, 2, 539-578, y 24, 1768, 111-180, pub. 1770 = *Œuvres*, 2, 581-652.

[36] 1776 = *Œuvres*, 4, 301-334.

[37] Ver la cita de Newton del cap. 17, sec. 3.

infinitas. De hecho, los términos «convergente» y «divergente» fueron empleados en 1668 por James Gregory, pero no desarrolló sus ideas. Newton reconoció la necesidad de examinar la convergencia, pero no pasó de afirmar que las series de potencias convergen para valores pequeños de la variable al menos tan bien como la serie geométrica; también señaló que algunas series pueden ser infinitas para algunos valores de x y por ello no tener utilidad, como, por ejemplo, la serie de $y = \sqrt{ax - x^2}$ en $x = a$.

También Leibniz mostró cierto interés acerca de la convergencia y en una carta del 25 de octubre de 1713 indicó a Jean Bernoulli lo que hoy es un teorema, a saber, que una serie cuyos términos alternan de signo y decrecen en valor absoluto monótamente hacia cero, es convergente [38].

Maclaurin, en su *Treatise of Fluxions* (1742), utilizó las series como un método sistemático de integración, afirmando que: «Cuando una fluente no puede representarse exactamente en términos algebraicos, ha de expresarse entonces mediante una serie convergente.» También admitió que los términos de una serie convergente han de decrecer continuamente y hacerse más pequeños que cualquier cantidad prescrita por pequeña que sea. «En ese caso, unos pocos términos del comienzo de la serie serán casi iguales al valor de la totalidad de aquéllos.» En el *Treatise*, Maclaurin dio el criterio integral (descubierto independientemente por Cauchy) para la convergencia de una serie: $\Sigma_n\ \phi(n)$ converge si y sólo si $\int_a^\infty \phi(x)$ es finita, supuesto que $\phi(x)$ es acotada y del mismo signo en $a \leq x \leq \infty$. Maclaurin lo dio en forma geométrica.

Nicolaus Bernoulli (1687-1759) también expuso algunas ideas sobre la convergencia en cartas a Leibniz de 1712 y 1713. En una carta del 7 de abril de 1713 [39], Bernoulli afirma que la serie

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2} x^2 + \cdots$$

no tiene suma cuando x es negativo y mayor en valor absoluto que 1 si n es fraccionario con denominador par. Es decir, la divergencia (aritmética) de una serie no es la única razón para que una serie no

[38] *Math. Schriften*, 3, 922-923. Leibniz dio también una demostración incorrecta en una carta a Jean del 10 de enero de 1714 = *Math. Schriften*, 3, 926.

[39] Leibniz: *Math. Schriften*, 3, 980-984.

tenga suma. Así, por ejemplo, si $x > 1$, las dos series

$$(1-x)^{-1/3} = 1 + \frac{1}{3}\,x + \frac{1\cdot 4}{3\cdot 6}\,x^2 + \frac{1\cdot 4\cdot 7}{3\cdot 6\cdot 9}x^3 + \cdots$$

$$(1-x)^{-1/2} = 1 + \frac{1}{2}\,x + \frac{1\cdot 3}{2\cdot 4}\,x^2 + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \cdots$$

son ambas divergentes, pero la primera tiene un valor posible mientras que la segunda tiene un valor imaginario. No se pueden distinguir porque los *restos están ausentes*. De todos modos, Nicolaus no estableció un concepto claro de convergencia. En una réplica del 28 de junio de 1713, Leibniz [40] utiliza el término «advergente» para series que convergen (más o menos en nuestro sentido) y está de acuerdo en que las series no advergentes pueden ser imposibles o infinitamente grandes.

No cabe duda de que Euler advirtió algunas de las dificultades que plantean las series divergentes, en particular, al utilizarlas para efectuar cálculos, pero desde luego no fue nada claro en lo referente a los conceptos de convergencia y divergencia. Sí advirtió que los términos han de hacerse infinitamente pequeños para que haya convergencia. Las cartas que se describen a continuación nos ilustran algo, indirectamente, sobre sus opiniones.

Nicolaus Bernoulli (1687-1759), en correspondencia con Euler durante 1742-43, había cuestionado algunas de las ideas y trabajos de éste. Señalaba que el uso de Euler en su artículo de 1734/35 (ver sec. 4) de

$$\operatorname{sen} s = s - \frac{s^3}{3!} + \frac{s^5}{5!} - \cdots = \left(1 - \frac{s^2}{\pi^2}\right)\left(1 - \frac{s^2}{4\pi^2}\right)\left(1 - \frac{s^2}{9\pi^2}\right)\cdots$$

implica efectivamente

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots,$$

pero que falta una demostración de la convergencia de la serie en s

[40] *Math. Schriften*, 3, 986.

de partida. En una carta del 6 de abril de 1743 [41], afirma que no puede imaginar que Euler crea que una serie divergente puede dar el valor exacto de una cantidad o de una función; señala que falta el resto. Así, $1/(1-x)$ no puede ser igual a $1+x+x^2+\cdots$ porque el resto, a saber, $x^{\infty-1}/(1-x)$, no está presente.

En otra carta de 1743, Bernoulli afirma que Euler debe distinguir entre una suma finita y una suma de infinitos términos; no hay último término en ese segundo caso y, por ello, no se puede aplicar a polinomios infinitos (como Euler hizo) la relación existente entre las raíces y los coeficientes de un polinomio de grado finito; para polinomios con un número infinito de términos no se puede hablar de la suma de las raíces.

No se conocen las respuestas de Euler a estas cartas de Bernoulli. En carta a Goldbach el 7 de agosto de 1745 [42], Euler se remite al argumento de Bernoulli de que series divergentes tales como

$$+1-2+6-24+120-720-\cdots$$

no poseen suma, pero afirma que estas series tienen un *valor* definido; señala que no se debe utilizar el término de «suma» porque éste se refiere a una adición efectiva y establece entonces lo que él entiende por un valor definido; explica que una serie divergente proviene de una expresión algebraica finita y afirma entonces que el valor de la serie es *el valor de la expresión algebraica de la que proviene la serie.* En el artículo de 1754/55 (sec. 4) añade: «Siempre que una serie infinita se obtenga como desarrollo de alguna expresión cerrada, puede utilizarse en operaciones matemáticas como equivalente de dicha expresión, incluso para los valores de la variable para los que la serie diverge.» Y repite el primer principio en su *Institutiones* de 1755:

Digamos, por tanto, que la suma de cualquier serie infinita es la expresión finita por cuyo desarrollo se genera la serie. En este sentido, la suma de la serie infinita $1-x+x^2-x^3+\cdots$ será $1/(1+x)$, porque la serie resulta como desarrollo de la fracción, cualquiera que sea el número que se pone en lugar de x. Si se conviene en esto, la nueva definición de la palabra suma coincide con el significado ordinario

[41] Fuss: *Correspondance*, 2, 701 y sigs.
[42] Fuss: *Correspondance*, 1, 324.

cuando una serie converge; y como las series divergentes no poseen suma en el sentido propio de la palabra, no surgen inconvenientes con esta terminología. Finalmente, por medio de esta definición, podemos conservar la utilidad de las series divergentes y defender su uso ante cualquier objeción [43].

Está bastante claro que Euler entendía esta doctrina limitada a las series de potencias.

En carta a Nicolaus Bernoulli en 1743, Euler admite que sí había tenido serias dudas en cuanto al uso de series divergentes, pero que nunca había sido inducido a error por el uso de su definición de suma [44]. A esto replicó Bernoulli que una misma serie podía resultar del desarrollo de dos funciones diferentes y en este caso la suma no sería única [45]. Después de esto, Euler escribió a Goldbach (en la carta del 7 de agosto de 1745) que «Bernoulli no da ejemplos y yo no creo que sea posible que la misma serie pueda provenir de dos expresiones algebraicas realmente distintas. De ahí que, indudablemente, cualquier serie, divergente o convergente, tiene una suma o valor definido».

Hubo una interesante continuación de este debate. Euler basaba en su argumento el que la suma de series tales como

$$1 - 1 + 1 - 1 + 1 \tag{58}$$

tomarían el valor de la función de la que proviene la serie. Así, la serie anterior proviene de $1/(1+x)$ cuando $x = 1$, por lo que su valor será 1/2. Sin embargo, unos cuarenta años más tarde, Jean-Charles (François) Callet (1744-99), en una memoria presentada a Lagrange (éste aprobó su publicación en las *Mémoires* de la Academia de Ciencias de París, pero la memoria nunca fue publicada), señaló que

$$\frac{1 + x + \cdots + x^{m-1}}{1 + x + \cdots + x^{n-1}} = \frac{1 - x^m}{1 - x^n} = 1 - x^m + x^n - x^{n+m} + x^{2n} - \cdots. \tag{59}$$

Entonces, para $x = 1$ (y $m < n$), como el primer miembro vale m/n,

---

[43] Parágrafos 108-111.

[44] *Opera Posthuma*, 1, 536.

[45] El 6 de abril de 1743; Fuss: *Correspondance*, 2, 701 y sigs.

la suma de la derecha ha de valer también m/n, donde m y n están a nuestra disposición.

Lagrange [46] examinó la objeción de Callet y argumentó que era incorrecta; utilizó el razonamiento probabilístico de Leibniz de la manera siguiente: supongamos $m = 3$ y $n = 5$. Entonces, la serie *completa* de la derecha en (59) es

$$1 + 0 + 0 - x^3 + 0 + x^5 + 0 + 0 - x^8 + 0 + x^{10} + 0 - \cdots.$$

Ahora, si se toma $x = 1$, la suma del primer término, de los dos primeros, de los tres primeros, . . ., entonces de cada cinco de esas sumas parciales, tres valen 1 y dos valen 0; en consecuencia, el valor más probable (valor medio) es 3/5; y este es el valor de la serie en (59) para $m = 3$ y $n = 5$. Por cierto que Poisson repite el argumento de Lagrange sin mencionar a éste [47].

Euler afirmó que había que proceder con mucha cautela en la suma de series divergentes; también distinguió entre series divergentes y series semiconvergentes, tales como (58), que oscilan en su valor cuando van añadiendo términos pero que no se hacen infinitas; llegó a conocer, por supuesto, la diferencia entre series convergentes y divergentes; en una ocasión (1747) en la que utilizó series para calcular la atracción que ejerce la tierra, en tanto que esferoide achatado, sobre una partícula situada en el polo, dice que la serie «converge impetuosamente».

También Lagrange fue consciente en cierta medida de la distinción entre convergencia y divergencia. En sus primeros trabajos fue sin duda descuidado en la materia; dice en un artículo [48] que una serie representará un número si converge a su extremidad, es decir, si su término n-ésimo se aproxima a 0. Más tarde, a finales del siglo XVIII, trabajando con la serie de Taylor, estableció lo que conocemos como teorema de Taylor [49], a saber,

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + \cdots + f^{(n)}(x)\frac{h^n}{n!} + R_n,$$

[46] *Mém. de l'Acad. des Sci., Inst. France*, 3, 1796, 1-11, pub. 1799; este artículo no aparece en las *Œuvres*.

[47] *Jour. de l'Ecole Poly.*, 12, 1823, 404-509. Si se impone utilizar la serie de *potencias* completa, entonces el argumento de Lagrange tiene más sentido. Se le puede dar rigor aplicando la definición de sumabilidad de Frobenius (cap. 47, sec. 4).

[48] *Hist. de l'Acad. de Berlin*, 24, 1770 = *Œuvres*, 3, 5-73, p. 61 en particular.

[49] *Théorie des fonctions*, 2.ª ed., 1813, cap. 6 = *Œuvres*, 9, 69-85. El teorema del

donde

$$R_n = f^{(n+1)}(x + \theta h) \frac{h^{n+1}}{(n+1)!},$$

valiendo θ entre 0 y 1. Esta expresión para R_n se conoce todavía como forma de Lagrange del resto. Lagrange indicó que la serie (finita) de Taylor no ha de de utilizarse sin tener en consideración el resto, pero no investigó la noción de convergencia ni la relación entre el valor del resto y la convergencia de la serie infinita; pensó que sólo era necesario considerar un número finito de términos de la serie, los suficientes para hacer el resto pequeño. La convergencia fue más tarde examinada por Cauchy, quien subrayó la importancia fundamental del teorema de Taylor así como el hecho de que, para obtener una serie convergente, el resto ha de tender a 0.

D'Alembert distinguió también las series convergentes de las divergentes. En su artículo «Série» de la *Encyclopédie* dice: «Cuando la progresión o la serie se aproxima cada vez más a una cantidad finita, y, consiguientemente, los términos de la serie, o cantidades de las que se compone, van disminuyendo, se dice que la serie es convergente, y si se continúa hasta infinito, se hará finalmente igual a dicha cantidad. Así, $1/2 + 1/4 + 1/8 + 1/16 + \cdots$ forma una serie que se aproxima constantemente a 1 y que se hará finalmente 1 cuando se continúe la serie hasta infinito.» En 1768, D'Alembert expresó sus dudas acerca del uso de series no convergentes; decía: «En cuanto a mí, confieso que todos los razonamientos basados en series que no son convergentes... me resultan muy sospechosos, incluso cuando los resultados concuerdan con verdades a las que se llega por otros caminos.» [50] En vista de las eficaces aplicaciones de las series llevadas a cabo por Jean Bernoulli y Euler, no fueron tenidas en cuenta, en el siglo XVIII, dudas como las expresadas por D'Alembert. En este mismo volumen, D'Alembert dio un criterio para la convergencia absoluta de la serie $u_1 + u_2 + u_3 + \cdots$, a saber, si para todo n mayor que un cierto r se tiene que $|u_{n+1}/u_n| < \varrho$ es independiente de n y menor que 1, la serie converge absolutamente [51].

valor medio del cálculo diferencial, $f(b) - f(a) = f'(c)(b - a)$, se debe a Lagrange (1797). Más tarde, se utilizó para deducir el teorema de Taylor como se hace en los libros modernos.

[50] *Opuscules mathématiques*, 5, 1768, 183.

[51] Páginas 171-182.

Edward Waring (1734-98), *Lucasian professor* de matemáticas en la universidad de Cambridge, sostuvo opiniones avanzadas sobre la convergencia; probó que

$$1+\frac{1}{2^n}+\frac{1}{3^n}+\frac{1}{4^n}+\cdots$$

converge cuando $n>1$ y diverge cuando $n<1$; estableció también el criterio bien conocido para la convergencia y la divergencia que hoy se denomina criterio del cociente y que se atribuye a Cauchy: se forma el cociente entre el término de lugar $(n+1)$ y el término de lugar n; si el límite de ese cociente cuando $n\to\infty$ es menor que 1, la serie converge, si es mayor que 1, la serie diverge y si es igual a 1 no se puede concluir nada.

Aunque Lacroix afirmó varias cosas absurdas sobre series en la edición de 1797 de su influyente *Traité du calcul différentiel et du calcul intégral*, fue más precavido en su segunda edición. Tratando de

$$\frac{a}{a-x}=1+\frac{x}{a}+\frac{x^2}{a^2}+\frac{x^3}{a^3}+\cdots,$$

dice que se debería hablar de la serie como un *desarrollo* de la función, ya que la serie no siempre tiene el *valor* de la función a la que corresponde [52]. La serie, dice, sólo da el valor de la función para $|x|<|a|$ y continúa con la idea ya expresada por Euler de que la serie está de todos modos asociada a la función para todo x; en cualquier trabajo analítico en que esté involucrada la serie será correcto concluir que se está tratando de hecho con la función; así pues, si descubrimos determinada propiedad de la serie, podemos estar seguros de que esa propiedad se verifica para la función; para darse cuenta de la certeza de este aserto, es suficiente observar que la serie verifica la ecuación que caracteriza a la función. Por ejemplo, para $y=a/(a-x)$ tenemos

$$a-(a-x)y=0.$$

Pero si se sutituye en esta ecuación y por la serie, veremos que la serie también la satisface; se sabe, prosigue Lacroix, que lo mismo

[52] 1810-19, 3 vols.; vol. 1, p. 4.

sería para cualquier otro ejemplo y remite al gran número de ellos que presenta en el texto.

Es justo decir que en los trabajos sobre series del siglo XVIII dominó el punto de vista formal. En general, los matemáticos incluso se resentían por cualquier limitación, tal como la necesidad de reflexionar acerca de la convergencia. Sus trabajos produían resultados útiles y quedaban satisfechos con tal pragmática sanción; sobrepasaron los límites de lo que podían justificar, pero al menos fueron prudentes en su utilización de las series divergentes. Como veremos, durante la mayor parte del siglo XIX predominó la insistencia en restringir el uso de series a las convergentes; pero, al final, los hombres del siglo XVIII fueron vindicados: dos ideas esenciales que ellos vislumbraron en las series infinitas habrían de ganar más tarde aceptación. La primera fue que las series divergentes pueden ser útiles para aproximaciones numéricas de funciones, y la segunda que una serie puede representar una función en operaciones analíticas, incluso aunque la serie sea divergente.

Bibliografía

Bernoulli, Jacques: *Ars Conjectandi*, 1713, reimpreso por Culture et Civilisation, 1968.

—: *Opera*, 2 vols., 1744. reimpresos por Birkhaüser, 1968.

Bernoulli, Jean: *Opera Omnia*, 4 vols., 1742, reimpresos por Georg Olms, 1968.

Burkhardt, H.: «Trigonometrische Reihen und Integrale bis etwa 1850», *Encyk. der math. Wiss.*, B. G. Teubner, 1914-15, 2, parte 1, pp. 825-1354.

—: «Entwicklungen nach oscillirenden Funktionen», *Jahres. der Deut. Math. Verein.*, vol. 10, 1908, pp. 1-1804.

—: «Über den Gebrauch divergenter Reihen in der Zeit: 1750-1860», *Math. Ann.*, 70, 1911, 189-206.

Cantor, Mortiz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teubner, 1898, vol. 3, caps. 85, 86, 97, 109 y 110.

Dehn, M. y E. D. Hellinger: «Certain Mathematical Achievements of James Gregory», *Amer. Math. Monthly*, 50, 1943, 149-163.

Euler, Leonhard: *Opera Omnia*, (1), vols. 10, 14 y 16 (2 partes), B. G. Teubner y Orell Füssli, 1913, 1924, 1933 y 1935.

Fuss, Paul Heinrich von: *Correspondance mathématique et physique de quel-*

ques célèbres géomètres du XVIIIème siècle, 2 vols., 1843, Johnson Reprint Corp., 1967.

Hofmann, Joseph E.: «Über Jakob Bernoullis Beiträge zur Infinitesimalmathematik», *L'Enseignement Mathématique*, (2), 2, 1956, 61-171; publicado también separadamente por el Institut de Mathématiques, Ginebra, 1957.

Montucla, J. F.: *Histoire des mathématiques*, A. Blanchard (reimpresión), 1960, vol. 3, pp. 206-243.

Reiff, R. A.: *Geschichte der unendlichen Reihen*, H. Lauppsche Buchhandlung, 1889; Martin Sändig (reimpresión), 1969.

Schneider, Ivo: «Der Mathematiker Abraham de Moivre: 1667-1754», *Archive for History of Exact Sciences*, 5, 1968, 177-317.

Smith, David Eugene: *A Source Book in Mathematics*, Dover (reimpresión), 1959, vol. 1, pp. 85-90 y 95-98.

Struik, D. J.: *A Source Book in Mathematics (1200-1800)*, Harvard University Press, 1969, pp. 111-115, 316-324, 328-333, 338-341 y 369-374.

Turnbull, H. W.: *James Gregory Tercetenary Memorial Volume*, Royal Society of Edinburgh, 1939.

—: *The Correspondence of Isaac Newton*, Cambridge University Press, 1959, vol. 1.

Capítulo 21
LAS ECUACIONES DIFERENCIALES ORDINARIAS EN EL SIGLO XVIII

Un viajero que rehúse pasar sobre un puente hasta haber comprobado personalmente la solidez de cada una de sus partes no irá muy lejos; es necesario arriesgar algo, incluso en matemáticas.

HORACE LLAMB

1. Las motivaciones

Los matemáticos trataron de aplicar el cálculo infinitesimal a resolver un número creciente de problemas físicos y pronto se vieron obligados a atacar una nueva clase de problemas. Lucharon más de lo que conscientemente se habían propuesto. Los problemas más simples conducían a cuadraturas que podían hallarse en términos de las funciones elementales y los algo más difíciles a cuadraturas que no podían expresarse de ese modo, como era el caso de las integrales elípticas (cap. 19, sec. 4), pero ambos tipos de problemas caían dentro de los límites del cálculo infinitesimal. Sin embargo, la resolución de problemas todavía más complicados requirió técnicas especializadas; así surgieron las ecuaciones diferenciales.

Fueron varias las clases de problemas físicos que motivaron la investigación en ecuaciones diferenciales. Una de ellas la constituían problemas en el campo de lo que hoy se conoce comúnmente como teoría de la elasticidad; un cuerpo es elástico si se deforma bajo la acción de una fuerza y recupera su forma original cuando se retira la fuerza. Los problemas de carácter más práctico tienen que ver con los perfiles que adoptan las vigas, verticales y horizontales, cuando

se les aplican cargas; estos problemas, tratados empíricamente por los constructores de las grandes catedrales medievales, fueron abordados matemáticamente por hombres tales como Galileo, Edme Mariotte (¿1620?-84), Robert Hooke (1635-1703) y Wren. El comportamiento de las vigas es una de las dos ciencias que Galileo trata en los *Diálogos relativos a dos nuevas ciencias*; la investigación de Hooke sobre muelles condujo a su descubrimiento de la ley que establece que la fuerza recuperadora de un muelle que se estira o se contrae es proporcional a la magnitud de la elongación o la contracción. Los pensadores del siglo XVIII, equipados con un nivel más alto en matemáticas, comenzaron sus trabajos en elasticidad abordando problemas como el del perfil que adopta un cable inelástico pero flexible suspendido de dos puntos fijos, el perfil de una cuerda, o una cadena, inelástica pero flexible suspendida de un punto fijo y que vibra, el perfil que adopta una cuerda que vibra fijada en sus extremos, el perfil de una barra cuando se fijan sus extremos y se la somete a una carga y el perfil de la barra cuando se pone a oscilar.

El péndulo continuó interesando a los matemáticos; la ecuación diferencial para el péndulo circular, $d^2\theta/dt^2 + mg$ sen $\theta = 0$, se resistía a ser tratada, e incluso quedaba por estudiar analíticamente la ecuación aproximada que se obtiene reemplazando sen θ por θ. Por otro lado, el período de un péndulo circular no es en rigor independiente de la amplitud, y se inició la búsqueda de una curva a lo largo de la cual se moviese el peso del péndulo de modo que el período fuese estrictamente independiente de la amplitud; Huygens había resuelto geométricamente este problema mediante la introducción de la cicloide, pero quedaba por elaborar la solución analítica.

El péndulo estaba íntimamente relacionado con otras dos investigaciones fundamentales del siglo XVIII, la forma de la Tierra y la verificación de la ley del cuadrado de los inversos de la atracción gravitatoria. El período aproximado de un péndulo, $T = 2\pi\sqrt{l/g}$, se utilizó para medir la fuerza de la gravedad en distintos puntos de la superficie de la Tierra, dado que el período depende de la aceleración g determinada por dicha fuerza; midiendo a lo largo de un meridiano longitudes sucesivas, correspondiendo cada longitud a un cambio de un grado en la latitud, se puede, con ayuda de alguna teoría y de los valores de g, determinar la forma de la Tierra; de hecho, Newton dedujo que la Tierra se abomba en el ecuador utilizando las variaciones observadas en el período en distintos lugares de la superficie terrestre.

Después de que Newton concluyese, mediante sus argumentos teóricos, que el radio ecuatorial era 1/230 más largo que el radio polar (valor que tiene un 30 por 100 de exceso), los científicos europeos estaban impacientes por confirmarlo. Un método consistiría en medir la longitud de un grado de latitud cerca del ecuador y cerca del polo; si la Tierra estuviese achatada, un grado de latitud sería ligeramente más largo en los polos que en el ecuador.

Jacques Cassini (1677-1756) y otros miembros de su familia realizaron tales medidas y llegaron en 1720 al resultado opuesto; encontraron que el diámetro de polo a polo era 1/95 más largo que el diámetro ecuatorial. Para resolver la cuestión, la Academia de Ciencias francesa envió en los años treinta del siglo dos expediciones, una a Laponia, bajo la dirección del matemático Pierre L. M. de Maupertuis, y la otra al Perú; el grupo de Maupertuis incluía un colega matemático, Alexis-Claude Clairaut. Sus mediciones confirmaron que la Tierra estaba achatada por los polos; Voltaire aclamó a Maupertuis como el «achatador de los polos y los Cassinis». De hecho, el valor de Maupertuis fue 1/178, menos preciso que el de Newton. La cuestión de la forma de la Tierra siguió siendo un tema de primera importancia y estuvo abierto por largo tiempo saber si la forma era un esferoide achatado, un esferoide alargado, un esferoide general o alguna otra figura de revolución.

El otro problema asociado, la verificación de la ley de la gravitación, se podría resolver si se conociese la forma de la Tierra. Dada la forma, se podría calcular la fuerza necesaria para mantener un objeto sobre o cerca de la superficie de la Tierra que gira; entonces, conociendo la aceleración g debida a la fuerza de la gravedad en la superficie, se puede comprobar si la fuerza de la gravedad total, que da lugar a la aceleración centrípeta y a g, responde realmente a la ley del cuadrado de los inversos. Clairaut, uno de los que cuestionaban la ley, creyó en un tiempo que debería ser de la forma $F = A/r^2 + B/r^3$. Ambos problemas de la ley de atracción y de la forma de la Tierra se entrelazan más adelante, pues cuando la Tierra se considera un fluido que gira en equilibrio, las condiciones para el equilibrio involucran la atracción mutua entre las partículas del fluido.

El campo de la física cuyo interés dominó el siglo fue la astronomía. Newton había resuelto lo que se conoce como el problema de los dos cuerpos, es decir, el movimiento de un único planeta bajo la atracción gravitatoria del Sol, idealizando cada uno de los cuerpos como una masa puntual. También había dado algunos pasos hacia

el tratamiento del problema fundamental de los tres cuerpos, el comportamiento de la Luna bajo la atracción de la Tierra y el Sol (cap. 17, sec. 3). Pero esto fue solamente el comienzo de los esfuerzos para estudiar los movimientos de los planetas y de sus satélites bajo la atracción gravitatoria del Sol, así como de la atracción mutua de todos los demás cuerpos. Incluso el trabajo de Newton en los *Principia*, aun consistiendo de hecho en la resolución de ecuaciones diferenciales, tenía que ser reformulado analíticamente, lo que se hizo gradualmente durante el siglo XVIII. Ello fue iniciado, por cierto, por Pierre Varignon, un buen físico y matemático francés, que trató de liberar a la dinámica del estorbo de la geometría. Newton ya había resuelto algunas ecuaciones diferenciales en forma analítica, por ejemplo, en su *Method of Fluxions* de 1671 (cap. 17, sec. 3); y en su *Tractatus* de 1676 hizo observar que la solución de $d^n y/dx^n = f(x)$ tiene un grado de arbitrariedad dado por un polinomio de grado $n - 1$. En la tercera edición de los *Principia*, proposición 34, escolio, se limita a un enunciado sobre qué formas de superficies de revolución ofrecen menor resistencia al movimiento en un fluido, pero en una carta a David Gregory de 1694 explica cómo logró sus resultados y en la explicación utiliza ecuaciones diferenciales.

Entre los problemas de astronomía, el que recibió mayor atención fue el del movimiento de la Luna, debido a que el método usual de determinar longitudes de los barcos en el mar (cap. 16, sec. 4), lo mismo que otros métodos recomendados en el siglo XVII, dependía del conocimiento en todo instante de la dirección de la Luna desde una posición fijada (que desde finales de siglo fue Greenwich, Inglaterra). Era necesario conocer esta dirección de la Luna con una precisión de 15 segundos de ángulo para determinar la hora en Greenwich con una precisión de un minuto; incluso ese margen de error podía llevar a un error de 30 kilómetros en la determinación de la posición del barco; pero con las tablas de la posición de la Luna disponibles en los tiempos de Newton, tal precisión estaba lejos de alcanzarse. Otra razón para el interés en la teoría del movimiento de la Luna es que puede utilizarse para predecir eclipses, lo que a su vez constituía una prueba de toda la teoría astronómica.

Las ecuaciones diferenciales ordinarias surgieron de los problemas que acabamos de esbozar. Al irse desarrollando la matemática, el campo de las ecuaciones en derivadas parciales dio lugar a investigaciones adicionales en ecuaciones diferenciales ordinarias, y lo mismo ocurrió con las ramas que hoy se conocen como geometría di-

ferencial y cálculo de variaciones. En este capítulo trataremos de los problemas que llevan directamente a los principales trabajos iniciales en ecuaciones diferenciales ordinarias, es decir, ecuaciones que contienen derivadas con respecto a una única variable independiente.

2. Ecuaciones diferenciales ordinarias de primer orden

Los primeros trabajos en ecuaciones diferenciales, como ocurrió con el cálculo infinitesimal a finales del siglo XVII y comienzos del XVIII, vieron primero la luz en cartas de un matemático a otro, muchas de las cuales ya no están disponibles, o en publicaciones que a menudo repiten los resultados establecidos o reivindicados en cartas. El anuncio de un resultado por uno de ellos provocaba frecuentemente la reivindicación por otro de que él había obtenido precisamente lo mismo con anterioridad, lo que, en vista de las encarnizadas rivalidades que existían, puede o no haber sido verdad. Algunas demostraciones se esbozaban meramente, y no está claro que los autores dispusiesen de su desarrollo completo; asimismo, supuestos métodos generales de solución simplemente se ilustraban mediante métodos particulares. Por estas razones, incluso aunque pasemos por alto todo lo relativo al rigor, es difícil atribuir los resultados obtenidos a la persona correcta.

En las *Acta Eruditorum* de 1693 [1] Huygens habla explícitamente de ecuaciones diferenciales, y Leibniz, en otro artículo de la misma revista y año [2] dice que las ecuaciones diferenciales son funciones de elementos del triángulo característico. Lo primero que normalmente aprendemos de las ecuaciones diferenciales, que aparecen al eliminar las constantes arbitrarias entre una función dada y sus derivadas, no se hizo hasta 1740, aproximadamente, y se debe a Alexis Fontaine del Bertins.

Jacques Bernoulli fue de los primeros en utilizar el cálculo infinitesimal para resolver problemas de ecuaciones diferenciales ordinarias; en mayo de 1690 [3] publicó su solución al problema de la isocrona, si bien Leibniz ya había dado una solución analítica; este problema consiste en encontrar una curva a lo largo de la cual un

[1] *Œuvres*, 10, 512-514.
[2] *Math. Schriften*, 5, 306.
[3] *Acta Erud.*, 1690, 217-219 = *Opera*, 1, 421-424.

péndulo tarde el mismo tiempo en efectuar una oscilación completa, sea grande o pequeño el arco que recorre; la ecuación, en los símbolos de Bernoulli, era

$$dy \sqrt{b^2 y - a^3} = dx \sqrt{a^3}.$$

Bernoulli concluía de la igualdad de diferenciales que las integrales (la palabra se utilizaba por primera vez) han de ser iguales y dio como solución

$$\frac{2b^2 y - 2a^3}{3b^2} \sqrt{b^2 y - a^3} = x \sqrt{a^3}.$$

La curva es, por supuesto, la cicloide.

En el mismo artículo de 1690, Jacques Bernoulli planteó el problema de encontrar la curva que adopta una cuerda flexible e inextensible colgada libremente de dos puntos fijos, la curva que Leibniz denominó catenaria. El problema había sido considerado ya en el siglo XV por Leonardo da Vinci; Galileo pensó que la curva era una parábola y Huygens afirmó que esto no era correcto, demostrando, básicamente por razonamientos físicos, que si el peso total de la cuerda y de las posibles cargas suspendidas de ella es uniforme por unidad horizontal, la curva es una parábola; para la catenaria, el peso por unidad es uniforme *a lo largo del cable.*

En las *Acta* de junio de 1691, Leibniz, Huygens y Jean Bernoulli publicaron soluciones independientes. La de Huygens era geométrica y confusa; Jean Bernoulli [4] la obtuvo por métodos del cálculo infinitesimal y la explicación completa se encuentra en su texto de cálculo infinitesimal de 1691; es la que ahora se expone en los textos de mecánica y de cálculo infinitesimal y está basada en la ecuación

$$\frac{dy}{dx} = \frac{s}{c}, \tag{1}$$

donde s es la longitud de arco entre B y un punto arbitrario A (fig. 21.1) y c depende del peso por unidad de longitud de la cuerda. Esta ecuación diferencial lleva a lo que ahora escribimos como $y = c \cosh (x/c)$. Leibniz también obtuvo este resultado por métodos del cálculo infinitesimal.

---

[4] *Acta Erud.*, 1691, 274-276 = *Opera*, 1, 48-51.

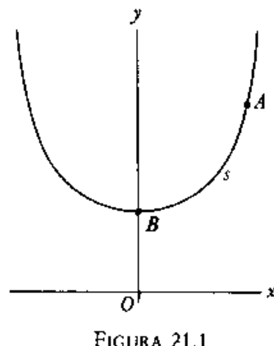


FIGURA 21.1

Jean Bernoulli se sintió inmensamente orgulloso de haber sido capaz de resolver el problema de la catenaria y que su hermano Jacques, que lo había propuesto, no lo hubiera conseguido. En una carta del 29 de septiembre de 1718 a Pierre Rémond de Montmort (1678-1719) se jacta de ello [5]:

Los esfuerzos de mi hermano no tuvieron éxito; en cuanto a mí, tuve más fortuna, ya que fui capaz (lo digo sin alarde, ¿por qué había de ocultar la verdad?) de resolverlo completamente y reducirlo a la rectificación de la parábola. Es cierto que me costó un esfuerzo que me robó el descanso por una noche entera; supuso mucho para aquellos días y para la escasa edad y práctica que yo tenía entonces, pero a la mañana siguiente, rebosante de alegría, corrí donde mi hermano, que estaba luchando tristemente con este, nudo gordiano sin llegar a ninguna parte, pensando todavía, como Galileo, que la catenaria era la parábola. ¡Alto! ¡Alto! le dije, no te tortures más intentando demostrar la identidad de la catenaria con la parábola, pues es completamente falsa; la parábola sirve efectivamente para la construcción de la catenaria, pero las dos curvas son tan diferentes que una es algebraica y la otra trascendente... Pero ahora me asombra usted afirmando que mi hermano descubrió un método para resolver este problema... Le pregunto, ¿cree usted realmente que si mi hermano hubiese resuelto el problema en cuestión hubiese sido tan considerado conmigo como para no figurar entre los que lo resolvieron, únicamente para cederme a mí la gloria de aparecer solo en el escenario en calidad de primero en resolverlo, junto con los señores Huygens y Leibniz?

---

[5] Johann Bernoulli, *Der Briefwechsel von Johann Bernoulli*, Birkhäuser Verlag, 1955, 97-98.

Durante los años 1691 y 1692 Jacques y Jean resolvieron también el problema del perfil que adopta una cuerda que cuelga y que sea flexible, inelástica y de densidad variable, el de una cuerda elástica de grosor constante y el de una cuerda sobre la que se aplica en cada uno de sus puntos una fuerza dirigida a un centro fijo; Jean resolvió también el problema inverso: dada la ecuación de la curva adoptada por una cuerda inelástica que cuelga, hallar la ley de la variación de la densidad de la cuerda con la longitud del arco. Las soluciones de Jean son las que aparecen con frecuencia en los textos de mecánica. Jacques publicó en las *Acta* de 1691 la demostración de que, de todos los perfiles que puede tomar una cuerda dada que cuelga de dos puntos fijos, la catenaria es la que posee el centro de gravedad más bajo.

En las *Acta* de 1691, Jacques Bernoulli dedujo la ecuación para la tractriz, la curva (fig. 21.2) para la que la razón de PT a OT es constante para cualquier punto P sobre la curva. Jacques obtuvo en primer lugar

$$\frac{dy}{ds} = \frac{y}{a},$$

donde s es la longitud de arco. De esta ecuación dedujo que

$$\int y\, dx = \int dy\, \sqrt{a^2 - y^2} \qquad (2)$$

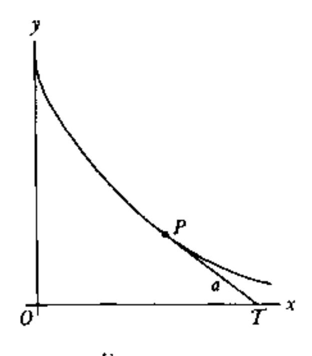


FIGURA 21.2

y

$$\int y^2\,dx = -\frac{1}{3}\sqrt{(a^2-y^2)^3}, \tag{3}$$

que él dejó como integrales que caracterizaban a la curva. (La ecuación (2) se puede integrar y da $x + \sqrt{a^2 - y^2} = a \log [(a + \sqrt{a^2 - y^2})/y]$.)

Leibniz descubrió la técnica de separación de variables y la comunicó a Huygens en una carta de 1691; resolvió así una ecuación de la forma $y(dx/dy) = f(x)\,g(y)$, escribiendo $dx/f(x) = g(y)\,dy/y$ y consiguiendo entonces integrar ambos miembros; no formuló el método general. También redujo (1691) la ecuación diferencial de primer orden homogénea, $y' = f(y/x)$, a cuadraturas, haciendo $y = vx$ y sustituyendo en la ecuación, lo que la convierte en una ecuación separable. Ambas ideas, separación de variables y solución de ecuaciones homogéneas, fueron expuestas más exhaustivamente por Jean Bernoulli en las *Acta Eruditorum* de 1694. A continuación, Leibniz, en 1694, probó cómo reducir la ecuación diferencial ordinaria de primer orden lineal $y' + P(x)y = Q(x)$ a cuadraturas, utilizando en su método un cambio en la variable dependiente; en general, Leibniz resolvió sólo ecuaciones diferenciales de primer orden.

Jacques Bernoulli propuso luego, en las *Acta* de 1695 [6], el problema de resolver la que hoy se conoce como ecuación de Bernoulli:

$$\frac{dy}{dx} = P(x)y + Q(x)y^n. \tag{4}$$

Leibniz demostró en 1696 [7] que se podía reducir a una ecuación lineal (primer grado en y e y') mediante el cambio de variable $z = y^{1-n}$. Jean Bernoulli dio otro método y Jacques la resolvió en las *Acta* de 1696, esencialmente por separación de variables.

En 1694, Leibniz y Jean Bernoulli introdujeron el problema de encontrar la curva o familia de curvas que cortan con un ángulo dado a una familia de curvas dada; Jean Bernoulli denominó trayectorias a las curvas secantes y señaló, partiendo del trabajo de Huygens sobre la luz, que este problema es importante para determinar las trayectorias de los rayos de luz que recorren un medio no uni-

---

[6] Página 553.

[7] *Acta Erud.*, 1696, 145.

forme porque dichos rayos cortan ortogonalmente los llamados frentes de onda de la luz. El problema no se hizo público hasta 1697, cuando Jean se lo planteó como un desafío a Jacques, quien lo resolvió en algunos casos especiales; Jean obtuvo la ecuación diferencial de las trayectorias ortogonales a una familia particular de curvas y la resolvió en 1698 [8]. Leibniz determinó las trayectorias ortogonales a una familia de curvas como sigue: consideremos $y^2 = 2bx$, donde b es el parámetro de la familia (un término que él introdujo); de esta ecuación se tiene $y\,dy/dx = b$; Leibniz hace entonces $b = -y\,dx/dy$, sustituye en $y^2 = 2bx$ y obtiene $y^2 = -2xy\,dx/dy$ como ecuación diferencial de las trayectorias; la solución es $a^2 - x^2 = y^2/2$. Aunque sólo resolvió casos especiales, Leibniz se hizo idea del problema general y del método para resolverlo.

El problema de las trayectorias ortogonales permaneció en estado latente hasta 1715, cuando Leibniz, apuntando ante todo a Newton, desafió a los matemáticos ingleses a descubrir el método general para encontrar las trayectorias ortogonales a una familia dada de curvas. Newton, cansado después de un día en la Casa de la Moneda, resolvió el problema antes de acostarse y la solución se publicó en las *Philosophical Transactions* de 1716 [9]. Newton también demostró cómo hallar las curvas que cortan a una familia dada con un ángulo constante o con un ángulo que varía con cada curva de la familia supuesta conforme a una ley dada; el método no es muy diferente del actual, aunque Newton utilizó ecuaciones de segundo orden.

También trabajó sobre este problema Nicolaus Bernoulli (1695-1726) en 1716. Jacob Hermann (1678-1733), un discípulo de Jacques Bernoulli, dio en las *Acta* de 1717 la regla de que si $F(x,y,c) = 0$ es la familia de curvas dada, entonces $y' = -F_x/F_y$, donde F_x y F_y son las derivadas parciales de F, y las trayectorias ortogonales tienen como pendiente F_y/F_x. [10] De aquí, afirmaba Hermann, resulta que la ecuación diferencial ordinaria de las trayectorias ortogonales a $F(x,y,c) = 0$ es

$$F_y dx = F_x dy. \qquad (5)$$

Despejaba c en (5), sustituía este valor en la ecuación original

[8] *Opera*, 1, 266.
[9] *Phil. Trans.*, 29, 1716, 399-400.
[10] *Acta Erud.*, 1717, 349 y sigs. También en Jean Bernoulli, *Opera*, 2, 275-279.

$F(x, y, c) = 0$ y resolvía la ecuación diferencial resultante. En realidad, este era el método de Leibniz, aunque establecido más explícitamente. Hoy se suele más bien hallar primero la ecuación diferencial que satisface $F = 0$, ecuación que ya no contiene el parámetro c, se reemplaza en ella y' por $-1/y'$ y se obtiene así la ecuación de las trayectorias ortogonales.

Jean Bernoulli planteó otros problemas de trayectorias a los ingleses, siendo Newton su pesadilla particular; como los ingleses y los continentales estaban ya enfrentados, los desafíos estaban marcados por el encarnizamiento y la hostilidad.

Jean Bernoulli resolvió después el problema de determinar el movimiento de un proyectil en un medio cuya resistencia es proporcional a una potencia de la velocidad; la ecuación diferencial es en este caso

$$m \frac{dv}{dt} - kv^n = mg. \qquad (6)$$

También fueron identificadas las ecuaciones de primer orden exactas, es decir, las ecuaciones $M(x, y)\, dx + N(x, y)\, dy = 0$, para las cuales $N\, dx + N\, dy$ es la diferencial exacta de una cierta función $z = f(x, y)$. Clairaut, célebre por su trabajo sobre la forma de la Tierra, había dado la condición $\partial M/\partial y = \partial N/\partial x$ para que la ecuación fuese exacta en sus artículos de 1739 y 1740 (cap. 19, sec. 6), condición que también fue dada independientemente por Euler en un artículo escrito en 1734-35 [11]. Si la ecuación es exacta, entonces, como señalaron Clairaut y Euler, se puede integrar.

Cuando una ecuación de primer orden no es exacta, es posible muchas veces multiplicarla por una cantidad, llamada factor integrante, que la convierte en exacta. Aunque se habían utilizado factores integrantes en problemas especiales, fue Euler quien se dio cuenta (en el artículo de 1734/35) de que este concepto proporcionaba un método de integración; estableció clases de funciones para las que existen factores integrantes y demostró también que si se conocen dos factores integrantes de una ecuación diferencial ordinaria de primer orden, entonces su cociente es una solución de la ecuación. Clairaut introdujo independientemente la idea de factor

[11] *Comm. Acad. Sci. Petrop.*, 7, 1734/35, 174-193, pub. 1740 = *Opera*, (1), 22, 36-56.

integrante en su artículo de 1739 y amplió la teoría en el artículo de 1740. Hacía 1740 se conocían todos los métodos elementales para resolver euaciones de primer orden.

3. Soluciones singulares

La soluciones singulares no se obtienen de la solución general dando un valor concreto a la constante de integración; es decir, no son soluciones particulares. Esto fue observado por Brook Taylor en su obra *Methodus Incrementorum* [12] al resolver una cuestión particular de primer orden y segundo grado. Leibniz había ya señalado en 1694 que una envolvente de una familia de soluciones es también solución. Las soluciones singulares fueron estudiadas más ampliamente por Clairaut y Euler.

El trabajo de Clairaut de 1734 [13] trata de la ecuación que ahora lleva su nombre,

$$y = xy' + f(y'). \tag{7}$$

Denotemos y' por p; entonces,

$$y = xp + f(p). \tag{8}$$

Derivando respecto a x, Clairaut obtenía

$$p = p + \{x + f'(p)\} \frac{dp}{dx}.$$

Entonces

$$\frac{dp}{dx} = 0 \quad \text{y} \quad x + f'(p) = 0. \tag{9}$$

La ecuación $dp/dx = 0$ da $y' = c$, por lo que de la ecuación original se tendrá

$$y = cx + f(c). \tag{10}$$

[12] 1715, p. 26.
[13] *Hist. de l'Acad. des Sci., Paris,* 1734, 196-215.

Esta es la solución general, consistente en una familia de rectas. El segundo factor, $x + f'(p) = 0$, se puede utilizar junto con la ecuación original para eliminar p, obteniéndose así una nueva solución, la solución singular. Para ver que se trata de la envolvente de la solución general tomamos (10) y derivamos respecto a c. Resulta

$$x + f'(c) = 0. \qquad (11)$$

La envolvente es la curva que resulta de eliminar c entre (10) y (11); pero estas dos ecuaciones son exactamente las mismas que las que proporcionan la solución singular. El hecho de que la solución singular sea una envolvente no llegó a verse entonces, pero Clairaut fue explícito en cuanto a que la solución singular no está incluida en la solución general.

Clairaut y Euler habían desarrollado un método para hallar la solución singular a partir de la propia ecuación, a saber, eliminando y' de $f(x,y,y') = 0$ y $\partial f/\partial y' = 0$. Este hecho y el de que las soluciones singulares no estén contenidas en la solución general intrigaron a Euler; en sus *Institutiones* de 1768 [14] dio un criterio para distinguir la solución singular de una integral particular que podía utilizarse cuando no se conocía la solución general, criterio que fue mejorado por D'Alembert [15]. A cotinuación, Laplace [16] extendió el concepto de soluciones singulares (que él llamó integrales particulares) a ecuaciones de orden superior y a ecuaciones diferenciales en tres variables.

Lagrange [17] llevó a cabo un estudio sistemático de las soluciones singulares y su relación con la solución general, dando el método general para obtener la solución singular a partir de la solución general por eliminación de la constante de una manera clara y elegante que supera la contribución de Laplace. Dada la solución general $V(x,y,\alpha) = 0$, el método de Lagrange consistía en encontrar $dy/d\alpha$, igualarla a 0 y eliminar α entre esta ecuación y $V = 0$; se puede utilizar el mismo procedimiento con $dx/d\alpha = 0$. También proporcionó información complementaria sobre el método de Clairaut y Euler para obtener la solución singular a partir de la ecuación diferencial

[14] Volumen 1, págs. 393 y sigs.

[15] *Hist. de l'Acad. des Sci., Paris*, 1769, 85 y sigs., pub. 1772.

[16] *Hist. de l'Acad. des Sci., Paris*, 1772, parte 1, 344 y sigs., pub. 1775 = *Œuvres*, 8, 325-366.

[17] *Nouv. Mém. de l'Acad. de Berlin*, 1774, pub. 1776 = *Œuvres*, 4, 5-108.

y, finalmente, dio la interpretación geométrica de la solución singular como envolvente de la familia de curvas integrales. Hubo un cierto número de dificultades especiales en la teoría de soluciones singulares que Lagrange no vio, como, por ejemplo, que pueden aparecer otras curvas singulares al eliminar y' entre $f(x,y,y') = 0$ y $\partial f/\partial y' = 0$ que no son soluciones singulares, o que una solución singular puede contener una rama que es una solución particular. La teoría detallada de las soluciones singulares fue desarrollada en el siglo XIX y fueron Cayley y Darboux quienes la expusieron en su forma actual en 1872.

4. Ecuaciones de segundo orden y ecuaciones de Riccati

Las ecuaciones diferenciales de segundo orden aparecen en problemas físicos ya en 1691. Jacques Bernoulli se planteó el problema de la forma de una vela bajo la presión del viento, el problema de la *velaria*, lo que le llevó a la ecuación de segundo orden $d^2x/ds^2 = (dy/ds)^3$, donde s es la longitud de arco. Jean Bernoulli trató este problema en su texto de cálculo infinitesimal de 1691 y estableció que es matemáticamente el mismo que el problema de la catenaria. Las ecuaciones de segundo orden aparecen a continuación al atacar el problema de determinar el perfil de una cuerda elástica que vibra, sujeta por los extremos —por ejemplo, una cuerda de violín—. Al abordar este problema, Taylor continuaba con un viejo tema; toda la cuestión de la matemática y los sonidos musicales comenzó con los pitagóricos, se prosiguió con ella en el período medieval y adquirió relevancia en el siglo XVII. Benedetti, Beeckman, Mersenne, Descartes, Huygens y Galileo destacan en esta materia, aunque no hubo nuevos resultados matemáticos que merezcan reseñarse aquí. El hecho de que una cuerda puede vibrar en muchos modos, esto es, en mitades, tercios, etc., y que el tono producido por una cuerda que vibra en k partes es el armónico k-ésimo (el fundamental es el primer armónico) era bien conocido en Inglaterra hacia 1700, en buena medida gracias a los trabajos experimentales de Joseph Sauveur (1653-1716).

Brook Taylor [18] obtuvo la frecuencia fundamental de una cuerda

[18] *Phil. Trans.*, 28, 1713, 26-32, pub. 1714; también en *Phil. Trans. Abridged*, 6, 1809, 7-12 y 14-17.

vibrante tensa; resolvió la ecuación $a\ddot{x}^2 = \dot{s}y\dot{y}$, donde $\dot{s} = \sqrt{\dot{x}^2 + \dot{y}^2}$ y la derivación es respecto al tiempo, y obtuvo $y = A$ sen (x/a) como perfil de la cuerda en todo instante. Aquí $a = l/\pi$, donde l es la longitud de la cuerda; el resultado de Taylor para la frecuencia fundamental (en notación moderna) es

$$v = \frac{1}{2l} \sqrt{T/\sigma},$$

donde T es la tensión de la cuerda, $\sigma = m/g$, m es la masa por unidad de longitud y g es la aceleración de la gravedad.

En sus esfuerzos para tratar la cuerda vibrante, Jean Bernoulli consideró, en una carta de 1727 a su hijo Daniel y en un artículo [19], una cuerda elástica sin peso cargada con n masas iguales e igualmente espaciadas; dedujo la frecuencia fundamental del sistema cuando hay 1, 2, . . ., 6 masas (hay otras frecuencias de oscilación del sistema de masas); Jean aplicó el que la fuerza sobre cada masa es $-K$ veces su desplazamiento, y resolvió $d^2x/dt^2 = -Kx$, integrando, por tanto, la ecuación del movimiento armónico simple por métodos analíticos; pasó después a la cuerda continua, de la que probó, como Taylor, que ha de tener el perfil de una curva sinusoidal (en todo instante) y calculó la frecuencia fundamental; resolvió, por tanto, la ecuación, $d^2y/dx^2 = -ky$. Ni Taylor ni Jean Bernoulli estudiaron los modos superiores de los cuerpos vibrantes elásticos.

Euler comenzó a considerar euaciones de segundo orden en 1728. Su interés en ellas fue suscitado en parte por sus trabajos en mecánica; había trabajado, por ejemplo, sobre el movimiento del péndulo en medios con rozamiento, lo que conduce a ecuaciones de segundo orden; trabajó para el rey de Prusia sobre el efecto de la resistencia del aire sobre los proyectiles; en esto, tomó el trabajo del inglés Benjamin Robins, lo mejoró y escribió una versión en alemán (1745), la cual fue traducida al francés y al inglés y utilizada en artillería.

Consideró también [20] una clase de ecuaciones de segundo orden que redujo mediante un cambio de variables a ecuaciones de primer orden. Consideró, por ejemplo, la ecuación

$$ax^m\, dx^p = y^n\, dy^{p-2}\, d^2y \qquad (12)$$

[19] *Comm. Acad. Sci. Petrop.*, 3, 1728, 13-28, pub. 1732 = *Opera*, 3, 198-210.

[20] *Comm. Acad. Sci. Petrop.*, 3, 1727, 124-137, pub. 1732 = *Opera*, (1), 22, 1-14.

o, en la forma de derivadas,

$$\left(\frac{dy}{dx}\right)^{p-2} \frac{d^2y}{dx^2} = \frac{ax^m}{y^n}. \tag{13}$$

Euler introdujo las nuevas variables t y v por medio de las ecuaciones

$$y = e^v t(v), \qquad x = e^{\alpha v}, \tag{14}$$

donde α es una constante a determinar. Las ecuaciones (14) se pueden contemplar como las ecuaciones paramétricas de x e y en términos de v, de modo que se pueden calcular dy/dx y d^2y/dx^2 y, sustituuendo en (13), obtener una ecuación de segundo orden en t como función de v. Euler fija entonces α de modo que quede eliminado el factor exponencial y v ya no aparezca explícitamente; una nueva transformación, a saber, $z = dv/dt$, reduce la ecuación de segundo orden a una de primer orden.

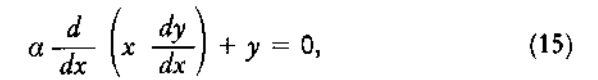
No merece la pena seguir con los detalles técnicos de este método porque se aplica únicamente a una clase de ecuaciones de segundo orden, pero históricamente este trabajo es significativo porque inicia el estudio sistemático de las ecuaciones de segundo orden y porque Euler introduce en él la función exponencial, que, como veremos, juega un papel importante en la resolución de las ecuaciones de segundo orden y de orden superior.

Antes de dejar San Petersburgo en 1733, Daniel Bernoulli terminó el artículo «Teoremas sobre oscilaciones de cuerpos conectados por un hilo flexible y de una cadena verticalmente suspendida» [21]. Comienza en él con la cadena suspendida de su extremo superior, sin peso pero cargada con pesos igualmente espaciados, y encuentra que, cuando la cadena se pone a vibrar, el sistema tiene distintos modos de (pequeña) oscilación alrededor de una línea vertical que pasa por el punto de suspensión; cada uno de estos modos tiene su propia frecuencia característica [22]. Entonces establece, para una cadena en suspensión que oscila, de densidad uniforme y de longitud

[21] *Comm. Acad. Sci. Petrop.*, 6, 1732/33, 108-122, pub. 1738.

[22] En el caso de n masas, cada masa tiene su propio movimiento, suma de n términos sinusoidales, teniendo cada uno de los cuales una de las frecuencias características. El sistema completo tiene n modos principales diferentes, cada uno de ellos con una de las frecuencias características; depende de las condiciones iniciales cuáles de éstas están presentes.

l, que el desplazamiento y a una distancia x del extremo inferior (fig. 21.3) satisface la ecuación



$$\alpha \frac{d}{dx}\left(x \frac{dy}{dx}\right) + y = 0, \qquad (15)$$



FIGURA 21.3

y que la solución es una serie infinita que, en notación moderna, se puede expresar como

$$y = AJ_0\,(2\sqrt{x/\alpha}), \qquad (16)$$

donde J_0 es la función de Bessel (de primera especie) de índice cero [23] Además, α es tal que

$$J_0\,(2\sqrt{l/\alpha}) = 0, \qquad (17)$$

siendo l la longitud de la cadena. Daniel Bernoulli afirma que (17) tiene infinitas raíces, que son decrecientes y se aproximan a 0, y da el valor máximo de α; para cada α hay un modo de oscilación y una frecuencia característica. Y dice en este punto: «Ni tampoco sería difícil derivar de esta teoría una teoría de cuerdas musicales que concuerde con la elaborada por Taylor y mi padre... Los experimentos muestran que en las cuerdas musicales hay intersecciones [nodos]

---

[23] $J_n(x) = \left(\frac{x}{2}\right)^n \sum_{k=0}^{\infty} \frac{(-1)^k\,(x/2)^{2k}}{k!\,(k+n)!}$ para n entero positivo o 0.

semejantes a los de las cadenas que vibran.» De hecho, va en esta cuestión más allá que Taylor y que su padre al identificar los modos o armónicos superiores de una cuerda vibrante.

Su artículo sobre la cadena en suspensión trata también de la cadena oscilante de grosor no uniforme, introduciendo en este caso la ecuación diferencial

$$\alpha \frac{d}{dx} \left(g(x) \frac{dy}{dx} \right) + y \frac{dg(x)}{dx} = 0, \tag{18}$$

donde $g(x)$ es la distribución de peso a lo largo de la cadena. Para $g(x) = x^2/l^2$, obtiene una solución en serie que, en notación moderna, se puede expresar en la forma:

$$y = 2A \left(\frac{2x}{\alpha} \right)^{-1/2} J_1(2\sqrt{2x/\alpha}) \tag{19}$$

con

$$J_1(2\sqrt{2l/\alpha}) = 0$$

y siendo J_1 la función de Bessel de primera especie de índice uno.

Lo que falta en las soluciones de Daniel Bernoulli es, en primer lugar, toda referencia al desplazamiento como función del *tiempo*, con lo que su trabajo queda, desde el punto de vista matemático, en el campo de las ecuaciones diferenciales ordinarias; y lo mismo, en segundo lugar, en cuanto a que los modos simples (los armónicos), que él identificó explícitamente como movimientos reales, se pueden superponer para formar movimientos más complicados.

Después de haber abordado el tema de los sonidos musicales en el libro *Tentamen Novae Theoriae Musicae ex Certissimis harmoniae Principiis Dilucide Expositae (Investigación sobre una nueva teoría de la música, claramente expuesta a partir de incontestables principios de la armonía)*, escrito antes de 1731 y publicado en 1739 [24], Euler prosiguió el trabajo de Daniel Bernoulli en el artículo «Sobre las oscilaciones de un hilo flexible cargado con una cantidad arbitraria de pesos» [25]. Los resultados de Euler son en buena parte los mismos de Bernoulli, salvo que los argumentos matemáticos de Euler son

---

[24] *Opera*, (3), 1, 197-427.

[25] *Comm. Acad. Sci. Petrop.*, 8, 1736, 30-47, pub. 1741 = *Opera*, (2), 10, 35-49.

más claros. Para una forma de cadena continua, a saber, el caso especial en que el peso es proporcional a x^n, Euler tuvo que resolver

$$\frac{x}{n+1}\,\frac{d^2y}{dx^2}+\frac{dy}{dx}+\frac{y}{\alpha}=0.$$

Obtiene la solución en serie que en notación moderna está dada por [26]

$$y=\mathrm{A}q^{-n/2}\,I_n(2\sqrt{q}),\qquad q=-\frac{(n+1)x}{\alpha}.$$

La n es aquí general, por lo que Euler está introduciendo funciones de Bessel de índice real arbitrario. Da también la solución en forma de integral

$$y=A\,\frac{\int_0^1(1-t^2)^{(2n-1)/2}\cosh\left(2t\sqrt{\dfrac{(n+1)x}{\alpha}}\right)dt}{\int_0^1(1-\tau^2)^{(2n-1)/2}\,d\tau}.$$

Este es quizá el primer caso de solución de una ecuación diferencial de segundo orden expresada como integral.

En un artículo de 1739 [27] Euler se ocupó de las ecuaciones diferenciales del oscilador armónico, $\ddot{x}+kx=0$, y de las oscilaciones forzadas del mismo

$$M\ddot{x}+Kx=F\ \mathrm{sen}\ \omega_a t. \qquad (20)$$

Obtuvo las soluciones por cuadraturas y descubrió (redescubrió, en realidad, ya que otros lo habían encontrado antes) el fenómeno de resonancia; a saber, que si ω es la frecuencia natural $\sqrt{K/M}$ del oscilador, que se obtiene cuando $F=0$, entonces, cuando ω_a/ω se aproxima a 1, la oscilación forzada tiene amplitud cada vez más grande y se hace infinita.

---

[26] Para v arbitrario (incluyendo valores complejos):

$$I_v(z)=\sum_{n=0}^{\infty}\frac{(z/2)^{v+2n}}{n!\,\Gamma(v+n+1)}.$$

Las $I_v(z)$ reciben el nombre de funciones de Bessel modificadas.

[27] *Comm. Acad. Sci. Petrop.*, 1, 1739, 128-149, pub. 1750 = *Opera*, (2), 10, 78-97.

En un intento de plantear un modelo para la transmisión del sonido en el aire, Euler considera en su artículo «Sobre la propagación de impulsos a través de un medio elástico» [28] n masas M conectadas por una especie de muelles (sin peso) que yacen en una recta horizontal PQ; se consideran movimientos longitudinales, es decir, a lo largo de PQ; teniéndose para la masa k-ésima

$$M\ddot{x}_k = K\,(x_{k+1} - 2x_k + x_{k-1}), \qquad k = 1, 2, \ldots, n$$

donde K es la constante del muelle y x_k es el desplazamiento de la masa k-ésima. Euler obtiene las frecuencias características correctas para cada una de las masas y la solución general

$$x_k = \sum_{r=1}^{n} A_r \operatorname{sen} \frac{rk\pi}{n+1} \cos\left(2\sqrt{K/M}\, t \quad \frac{\operatorname{sen} r \cdot \pi/2}{n+1}\right) \qquad (21)$$

$k = 1, 2, \ldots, n$. Así pues, no sólo obtiene los modos individuales para cada una de las masas sino también el movimiento de esa masa como suma de modos armónicos simples; cada modo particular que pueda aparecer depende de las condiciones iniciales, es decir, de cómo las masas se ponen en movimiento. Todos estos resultados se pueden interpretar en términos del movimiento transversal (perpendicular a PQ) de la cuerda cargada.

Alguna de las ecuaciones ya consideradas, como, por ejemplo, la ecuación de Bernoulli, no es lineal; es decir, en tanto que ecuación en las variables y, y' y y'' (si aparece), contiene términos de grado dos o superior. Entre las ecuaciones de primer orden de este tipo hay algunas que tienen un interés especial por estar relacionadas íntimamente con ecuaciones de segundo orden lineales. En los primeros tiempos de las ecuaciones diferenciales ordinarias recibió una gran atención la ecuación no lineal de Riccati.

$$\frac{dy}{dx} = a_0(x) + a_1(x)y + a_2(x)\,y^2. \qquad (22)$$

La ecuación de Riccati cobró importancia cuando Jacopo Fran-

[28] *Novi Comm. Acad. Sci. Petrop.*, 1, 1747/48, 67-105, pub. 1750 = *Opera*, (2), 10, 98-131.

cesco, conde Riccati de Venecia (1676-1754), quien trabajó en acústica, la introdujo para facilitar la resolución de ecuaciones diferenciales de segundo orden. Riccati consideró curvas cuyos radios de curvatura dependen sólo de las ordenadas, lo que le llevó [29] a

$$x^m \frac{d^2x}{dp^2} = \frac{d^2y}{dp^2} + \left(\frac{dy}{dp}\right)^2$$

(él escribió $x^m\, d^2x = d^2y + (dy)^2$), en donde se ha de entender que x e y dependen de p. Mediante cambios de variables obtuvo

$$x^m \frac{dq}{dx} = \frac{du}{dx} + \frac{u^2}{q},$$

que es una ecuación de primer orden. Riccati supuso después que q es una potencia de x, por ejemplo, x^n, para llegar a la forma

$$\frac{du}{dx} + \frac{u^2}{x^n} = nx^{m+n-1}, \qquad (23)$$

y demostró a continuación cómo resolver (23) para valores especiales de n por el método de separación de variables para ecuaciones diferenciales ordinarias. Más tarde, varios de los Bernoulli determinaron otros valores de n para los que se puede resolver (23) por separación de variables.

El trabajo de Riccati no es sólo importante por tratar ecuaciones de segundo orden sino por la idea de reducir ecuaciones de segundo orden al primero; la idea de reducir el orden de una ecuación diferencial por uno u otro medio llegaría a ser considerada como uno de los métodos principales para el tratamiento de ecuaciones diferenciales ordinarias de orden superior.

En 1760, Euler [30] consideró la ecuación de Riccati

$$\frac{dz}{dx} + z^2 = ax^n \qquad (24)$$

---

[29] *Acta Erud.*, 1724, 66-73.

[30] *Novi Comm. Acad. Sci. Petrop.*, 9, 1760/61, 3-63, pub. 1763 = *Opera*, (1), 22, 334-394, y 9, 1762-/63, 154-169, pub. 1764 = *Opera*, (1), 22, 403-420.

y demostró que si se conoce una integral particular v, entonces la transformación

$$z = v + u^{-1}$$

convierte a aquélla en una ecuación lineal. Además, si se conocen dos integrales particulares, la integración de la ecuación original se reduce a cuadraturas.

D'Alembert [31] fue el primero en considerar la forma general (22) de la ecuación de Riccati y en utilizar el término «ecuación de Riccati» para esta forma. Comenzó con

$$\frac{d^2S}{dx^2} = \frac{-\lambda^2 x\pi^2 S}{2aLe} \tag{25}$$

e hizo

$$S = \exp\left[\int p\,dx\right],\ p = f(x), \tag{26}$$

obteniendo así la forma (22) para una ecuación en p como función de x.

5. Ecuaciones de orden superior

En diciembre de 1734, Daniel Bernoulli escribió a Euler, quien estaba en San Petersburgo, que había resuelto el problema del desplazamiento transversal de una barra elástica (un cuerpo unidimensional de madera o de acero) fijado a una pared en uno de sus extremos y libre en el otro. Bernoulli había obtenido la ecuación diferencial

$$K^4 \frac{d^4y}{dx^4} = y, \tag{27}$$

donde K es una constante, x es la distancia desde el extremo libre de la barra e y es el desplazamiento vertical en ese punto respecto

[31] *Hist. de l'Acad. de Berlin*, 19, 1763, 242 y sigs., pub. 1770.

a la posición sin pandeo de la barra. Euler, en una réplica escrita antes de junio de 1735, afirmó que también él había descubierto esta ecuación y que no era capaz de integrarla salvo utilizando series, y que había obtenido cuatro series distintas; estas series representaban funciones circulares y exponenciales, pero Euler no lo vio entonces.

Cuatro años más tarde, en una carta a Jean Bernoulli (15 de septiembre de 1739), Euler indicaba que su solución se podía representar como

$$y = A\left[\left(\cos\frac{x}{K} + \cosh\frac{x}{K}\right) - \frac{1}{b}\left(\operatorname{sen}\frac{x}{K} + \operatorname{senh}\frac{x}{K}\right)\right], \quad (28)$$

donde b está determinado por la condición $y = 0$ cuando $x = l$, de modo que

$$b = \frac{\operatorname{sen}\dfrac{l}{K} + \operatorname{senh}\dfrac{l}{K}}{\cos\dfrac{l}{K} + \cosh\dfrac{l}{K}}.$$

Los problemas de elasticidad condujeron a Euler a considerar el problema matemático de la resolución de ecuaciones lineales generales con coeficientes constantes, y en una carta a Jean Bernoulli de 15 de septiembre de 1739 escribe que había tenido éxito. Bernoulli le respondió afirmando que él ya había considerado tales ecuaciones en 1700, incluso con coeficientes variables; en realidad, únicamente había considerado una ecuación especial de tercer orden de la que había probado que se podía reducir a una de segundo orden.

En la publicación de su trabajo [32], Euler considera la ecuación

$$0 = Ay + B\frac{dy}{dx} + C\frac{d^2y}{dx^2} + D\frac{d^3y}{dx^3} + \cdots + L\frac{d^ny}{dx^n}, \quad (29)$$

donde los coeficientes son constantes; la ecuación se dice que es homogénea porque el término independiente de y y sus derivadas es 0. Euler indica que la solución general ha de contener n constantes arbitrarias y que dicha solución vendrá dada por la suma de n so-

[32] *Misc. Berolin.*, 7, 1743, 193-242 = *Opera*, (1), 22, 108-149.

luciones particulares, cada una de ellas multiplicada por una constante; hace entonces la sustitución

$$y = \exp\left[\int r \ dx\right],$$

con r constante, y obtiene la ecuación en r

$$A + Br + Cr^2 + \cdots + Lr^n = 0,$$

que se denomina ecuación característica o auxiliar. Cuando q es una raíz real simple de esta ecuación, entonces

$$a \ \exp\left[\int q \ dx\right]$$

es una solución de la ecuación diferencial original. Si la ecuación característica tiene una raíz múltiple q, Euler hace $y = e^{qx}u(x)$ y sustituye en la ecuación diferencial, obteniendo que

$$y = e^{qx}(\alpha + \beta x + \gamma x^2 + \cdots + \varkappa x^{k-1}) \qquad (30)$$

es una solución que contiene k constantes arbitrarias si q aparece k veces como raíz de la ecuación característica. Trata también los casos de raíces complejas conjugadas y de raíces complejas múltiples, con lo que Euler resuelve completamente las ecuaciones lineales homogéneas con coeficientes constantes.

Algo más tarde [33] estudió la ecuación diferencial ordinaria lineal de orden n no homogénea; su método consistió en multiplicar la ecuación por $e^{\alpha x}$, integrar ambos miembros y proceder a determinar α de modo que la ecuación se reduzca a una de orden inferior. Así, por ejemplo, para resolver

$$C\frac{d^2y}{dx^2} + B\frac{dy}{dx} + Ay = X(x), \qquad (31)$$

multiplica todo por $e^{\alpha x}\,dx$ y obtiene

$$\int\left[e^{\alpha x}C\frac{d^2y}{dx^2} + e^{\alpha x}B\frac{dy}{dx} + e^{\alpha x}Ax\right]dx = \int e^{\alpha x}X\,dx.$$

[33] *Novi Comm. Acad. Sci. Petrop.*, 3, 1750/51, 3-35, pub. 1753 = *Opera*, (1), 22, 181-213.

Pero, para A', b' y α apropiados, el primer miembro ha de ser

$$e^{\alpha x}\left(A'y + B' \frac{dy}{dx}\right).$$

Derivando esta expresión y comparando con la ecuación original, Euler obtiene que

$$B' = C, \qquad \mathrm{A}' = B - \alpha C, \qquad \mathrm{A}' = A/\alpha \tag{32}$$

con lo cual, de las dos últimas ecuaciones,

$$A - B\alpha + C\alpha^2 = 0. \tag{33}$$

Quedan, pues, determinados α, A' y B' y la ecuación original se reduce a

$$\mathrm{A}'y + B' \frac{dy}{dx} = e^{-\alpha x} \int e^{\alpha x} X \, dx. \tag{34}$$

Un factor integrante de esta ecuación es $e^{\beta x}\,dx$, donde $\beta = A'/B'$, de modo que, por (32), se tiene $\alpha\beta = A/C$ y $\alpha + \beta = B/C$ y consiguientemente, por (33), α y β son las raíces de $A - B\alpha + C\alpha^2 = 0$.

El método se aplica a ecuaciones diferenciales ordinarias de orden n lineales, reduciendo el orden paso a paso como en el ejemplo anterior. Euler también consideró los casos de raíces iguales de la ecuación en α y de raíces complejas.

Siguiendo las líneas del tratamiento de las ecuaciones lineales con coeficientes constantes, Lagrange pasó a considerar el caso de coeficientes variables [34]. Esto lleva, como veremos, al concepto de ecuación adjunta; Lagrange comienza con

$$Ly + M \frac{dy}{dt} + N \frac{d^2y}{dt^2} + \cdots = T, \tag{35}$$

donde L, M, N y T son funciones de t. (Nos limitamos, para simplificar, a ecuaciones de segundo orden.) Lagrange multiplica por

[34] *Misc. Taur.*, 3, 1762/65, 179-186 = *Œuvres*, 1, 471-478.

$z\,dt$, donde $z(t)$ es una función a determinar, e integra por partes:

$$\int Mzy'\,dt = Mzy - \int (Mz)'y\,dt$$

$$\int Nzy''\,dt = Nzy' - (Nz)'y + \int (Nz)''y\,dt.$$

Entonces, la ecuación original se convierte en

$$y[Mz - (Nz)'] + y'(Nz) + \int [Lz - (Mz)' + (Nz)'']y\,dt = \int Tz\,dt.$$

El corchete bajo el signo integral se puede tratar como una ecuación diferencial ordinaria en z igualándolo a 0; si se puede resolver en $z(t)$, quedará para y una ecuación diferencial ordinaria de orden inferior al original. La nueva ecuación diferencial en z se dice que es la adjunta de la ecuación original, el término fue introducido por Lazare Fuchs en 1873. Lagrange no utilizó ningún término especial para ella.

Para tratar la ecuación en z (la ecuación adjunta), Lagrange procede del mismo modo que para reducir el orden. Multiplica por $w(t)dt$, hace como antes y llega a una ecuación en w que permite reducir el orden de la ecuación en z; la ecuación en w resulta ser la ecuación original (35), excepto que el segundo miembro es 0. Así pues, Lagrange descubrió el teorema que dice que la adjunta de la adjunta de la ecuación diferencial ordinaria no homogénea original es la ecuación homogénea asociada a la original. Euler hizo esencialmente lo mismo en 1778; había conocido el trabajo de Lagrange, pero al parecer lo olvidó.

En trabajos posteriores sobre ecuaciones homogéneas con coeficientes variables, Lagrange [35] extendió a estas ecuaciones algunos de los resultados que Euler había obtenido para las ecuaciones con coeficientes constantes. Lagrange descubrió que la solución general de la ecuación homogénea es una suma de soluciones particulares independientes, cada una de ellas multiplicada por una constante arbitraria, y que conociendo m integrales particulares de una ecuación homogénea de orden n se puede reducir el orden en m unidades.

---

[35] *Misc. Taur.*, 3, 1762/65, 190-199 = *Œuvres*, 1, 481-490.

6. El método de integración por series

Ya hemos tenido ocasión de señalar que algunas ecuaciones diferenciales se resolvían por medio de series. La importancia de este método, incluso en la actualidad, justifica el dedicar unos pocos comentarios específicos a la materia. Las soluciones en serie han sido utilizadas tan ampliamente desde 1700 que hemos de limitarnos a unos pocos ejemplos.

Sabemos que Newton utilizó series para integrar funciones algo complicadas, incluso cuando sólo estaban implicadas cuadraturas. También las utilizó para resolver ecuaciones de primer orden; así, para integrar

$$\dot{y} = 2 + 3x - 2y + x^2 + x^2 y, \tag{36}$$

Newton supone que

$$y = A_0 + A_1 x + A_2 x^2 + \cdots. \tag{37}$$

Entonces

$$\dot{y} = A_1 + 2A_2 x + 3A_3 x^2 + \cdots. \tag{38}$$

Sustituyendo (37) y (38) en (36) e igualando coeficientes de las distintas potencias de x, resulta

$$A_1 = 2 - 2A_0, \qquad 2A_2 = 3 - 2A_1, \qquad 3A_3 = 1 + A_0 - 2A_2 + \cdots$$

Se determinan así los A_i; el hecho de que A_0 queda indeterminado y que por lo tanto existen infinitas soluciones sí fue advertido, pero la importancia de una constante arbitraria no fue plenamente apreciada hasta aproximadamente 1750. Leibniz resolvió algunas ecuaciones diferenciales elementales por medio de series [36] y utilizó también el método anterior de coeficientes indeterminados.

Euler puso en un primer plano el método de integración por series desde aproximadamente 1750 en adelante a fin de resolver ecuaciones diferenciales que no podían integrarse en forma cerrada. Aunque él trabajó con ecuaciones diferenciales concretas y los de-

---

[36] *Acta Erud.*, 1693 = *Math. Schriften*, 5, 285-288.

talles de lo que hizo son a menudo complicados, su método es el que utilizamos en la actualidad. Euler supone que se tiene una solución de la forma

$$y = x^{\lambda}(A + Bx + Cx^2 + \cdots),$$

sustituye y y sus derivadas en la ecuación diferencial y determina λ y los coeficientes A, B, C, ... a partir de la condición de que sea cero el coeficiente de cada una de las potencias de x en la serie resultante. Así, la ecuación que aparece en su trabajo sobre la membrana oscilante [37] (ver cap. 22, sec, 3), a saber,

$$\frac{d^2u}{dr^2} + \frac{1}{r}\frac{du}{dr} + \left(\alpha^2 - \frac{\beta^2}{r^2}\right)u = 0,$$

hoy llamada ecuación de Bessel; Euler la resolvió por medio de una serie, dando la solución

$$u(r) = r^{\beta}\left\{1 - \frac{1}{1\cdot(\beta+1)}\left(\frac{\alpha r}{2}\right)^2 + \frac{1}{1\cdot 2(\beta+1)(\beta+2)}\left(\frac{\alpha r}{2}\right)^4 - \frac{1}{1\cdot 2\cdot 3(\beta+1)(\beta+2)(\beta+4)}\left(\frac{\alpha r}{2}\right)^6 + \cdots\right\},$$

que es, salvo un factor que depende sólo de β, lo que ahora escribimos como $J_\beta(r)$. En un trabajo posterior sobre estas funciones, demostró que, para valores de β semienteros, la serie se reduce a funciones elementales; observó además que $u(r)$, para β real, tiene un número infinito de ceros y dio una representación integral para $u(r)$. Finalmente, Euler dio para $\beta = 0$ y $\beta = 1$ la segunda solución (en serie) independiente de la ecuación diferencial.

En las *Institutiones Calculi Integralis* [38], Euler trató la ecuación diferencial hipergeométrica.

$$x(1-x)\frac{d^2y}{dx^2} + [c - (a+b+1)x]\frac{dy}{dx} - aby = 0 \qquad (39)$$

[37] *Novi Comm. Acad. Sci. Petrop.*, 10, 1764, 243-260, pub. 1766 = *Opera*, (2), 10, 344-359.

[38] Volumen 2, 1769, caps. 8-11.

de la que dio la solución en serie

$$y = 1 + \frac{a \cdot b}{1 \cdot c}\, x + \frac{a(a+1)\, b(b+1)}{1 \cdot 2 \cdot c\,(c+1)}\, x^2$$

$$+ \frac{a(a+1)\,(a+2)\, b(b+1)\,(b+2)}{1 \cdot 2 \cdot 3 \cdot c\,(c+1)\,(c+2)}\, x^3 + \cdots \qquad (40)$$

En su principal artículo sobre la materia, escrito en 1778 [39], volvió a considerar la ecuación (39) y la solución (40). Había escrito otros artículos sobre lo que él llamaba serie hipergeométrica, pero en ellos el término se refería a otra serie originalmente introducida por Wallis; el término «hipergeométrica», para describir la ecuación diferencial (39) y la serie (40), se debe a Johann Friedrich Pfaff (1765-1825), amigo y maestro de Gauss. La serie (40) se denota en la actualidad por $F(a,b,c;z)$. Con esta notación, Euler probó las célebres relaciones

$$F(-n, b, c; z) = (1-z)^{c+n-b}\, F(c+n, c-b, c; z)$$

$$F(-n, b, c; z) = \frac{n!}{c(c+1) \cdots (c+n+1)}. \qquad (41)$$

$$\int_0^1 t^{-n-1}(1-t)^{c+n-1}\,(1-tz)^{-b}\, dt.$$

7. Sistemas de ecuaciones diferenciales

Las ecuaciones diferenciales implicadas hasta aquí en el estudio de la elasticidad eran bastante simples, debido a que los matemáticos utilizaban principios físicos un tanto toscos mientras continuaban esforzándose en captar otros más refinados. En el campo de la astronomía, sin embargo, los principios físicos, fundamentalmente las leyes del movimiento de Newton y la ley de gravitación, estaban claros y los problemas matemáticos eran mucho más profundos. El problema matemático fundamental al estudiar el movimiento de dos

[39] *Nova Acta Acad. Sci. Petrop.*, 12, 1794, 58-70, pub. 1801 = *Opera*, (1), 16_2, 41-55.

o más cuerpos, moviéndose cada uno bajo la atracción gravitatoria de los otros, es el de resolver un sistema de ecuaciones diferenciales ordinarias, aunque el problema se reduce con frecuencia a la resolución de una única ecuación.

Aparte de casos aislados, los trabajos sobre sistemas de ecuaciones se refirieron principalmente a problemas de astronomía. El punto de partida para escribir las ecuaciones diferenciales es la segunda ley del movimiento de Newton, $f = ma$, donde f es la fuerza de atracción; se trata de una ley vectorial, lo que significa que cada componente de f produce una aceleración en la dirección de la componente. En un artículo de 1750 [40], Euler expresó analíticamente la segunda ley de Newton en la forma

$$f_x = m\frac{d^2x}{dt^2}, \qquad f_y = m\frac{d^2y}{dt^2}, \qquad f_z = m\frac{d^2z}{dt^2}. \tag{42}$$

Supone aquí ejes rectangulares fijados y señala también que para cuerpos puntuales, es decir, cuerpos que se pueden considerar como si sus masas estuviesen concentradas en un punto, m es la masa total, mientras que para cuerpos distribuidos, m es dM.

Consideraremos brevemente las correspondientes ecuaciones diferenciales. Supongamos que un cuerpo fijo de masa M está situado en el origen y que un cuerpo móvil de masa m está en (x, y, z). Entonces, las componentes de la fuerza gravitatoria en las direcciones de los ejes (fig. 21.4) son

$$f_x = -\frac{GMmx}{r^3}, \qquad f_y = -\frac{GMmy}{r^3}, \qquad f_z = -\frac{GMmz}{r^3},$$

donde G es la constante de gravitación y $r = \sqrt{(x^2 + y^2 + z^2)}$. Se puede ver fácilmente que el cuerpo móvil permanece en un plano de manera que el sistema de ecuaciones (42) se reduce a

$$\frac{d^2x}{dt^2} = -\frac{kx}{r^3}, \qquad \frac{d^2y}{dt^2} = -\frac{ky}{r^3} \tag{43}$$

---

[40] *Hist. de l'Acad. de Berlin*, 6, 1750, 185-217, pub. 1752 = *Opera*, (2), 5, 81-108.

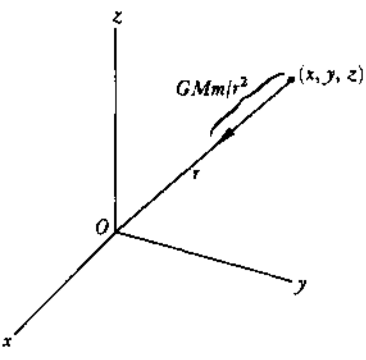


FIGURA 21.4

con $k = GM$. En coordenadas polares estas ecuaciones se convierten en

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{k}{r^2}$$

$$r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} = 0. \qquad (44)$$

En este caso de un cuerpo que se mueve bajo la fuerza de atracción de otro que permanece fijo, las dos ecuaciones diferenciales se pueden combinar para dar una sola que contiene x e y o r y θ, ya que, por ejemplo, la segunda ecuación polar se puede integrar para dar $r^2\, d\theta/dt = C$, y el valor de $d\theta/dt$ se puede sustituir en la primera ecuación. Se deduce que el cuerpo móvil describe una sección cónica con un foco en la posición que ocupa el otro cuerpo.

Si los dos cuerpos se mueven, cada uno sometido a la atracción del otro, las ecuaciones diferenciales son algo diferentes. Sean m_1 y m_2 las masas de dos cuerpos esféricos de masa con simetría esférica, tales que $m_1 + m_2 = M$. Elijamos un sistema de coordenadas fijo (normalmente se toma como origen el centro de gravedad de los dos cuerpos) y sean (x_1, y_1, z_1) las coordenadas de un cuerpo y (x_2, y_2, z_2) las coordenadas del otro; sea $r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$ la distancia entre ambos. Entonces, el sistema de ecuaciones que

describe el movimiento es

$$m_1 \frac{d^2x_1}{dt^2} = -km_1 m_2 \frac{(x_1 - x_2)}{r^3}, \qquad m_1 \frac{d^2y_1}{dt^2} = -km_1 m_2 \frac{(y_1 - y_2)}{r^3},$$

$$m_1 \frac{d^1z_1}{dt^2} = -km_1 m_2 \frac{(z_1 - z_2)}{r^3}, \qquad m_2 \frac{d^2x_2}{dt^2} = -km_1 m_2 \frac{(x_2 - x_1)}{r^3},$$

$$m_2 \frac{d^2y_2}{dt^2} = -km_1 m_2 \frac{(y_2 - y_1)}{r^3}, \qquad m_2 \frac{d^2z_2}{dt^2} = -km_1 m_2 \frac{(z_2 - z_1)}{r^3}.$$

Este es un sistema de seis ecuaciones de segundo orden cuya solución requiere doce integrales, cada una con una constante arbitraria de integración; estas constantes quedan determinadas por las tres coordenadas de la posición inicial y las tres componentes de la velocidad inicial de ambos cuerpos. Las ecuaciones pueden resolverse y se comprueba que cada uno de los cuerpos se mueve en una sección cónica respecto al centro de gravedad común.

De hecho, este problema del movimiento de dos esferas bajo la atracción mutua debida a la fuerza gravitatoria fue resuelto geométricamente por Newton en los *Principia* (libro I, sección 11). Sin embargo, no se emprendieron trabajos analíticos por algún tiempo. En mecánica, los franceses seguían el sistema de Descartes hasta que Voltaire, después de visitar Londres en 1727, regresó apoyando a Newton. Incluso Cambridge, la propia universidad de Newton, continuaba enseñando filosofía natural por el texto de Jacques Rohault (1620-75), un cartesiano. Además, los matemáticos más eminentes de finales del siglo XVII —Huygens, Leibniz y Jean Bernoulli— se oponían al concepto de gravedad y, en consecuencia, a sus aplicaciones. Los métodos analíticos para tratar el movimiento de los planetas fueron abordados por Daniel Bernoulli, quien recibió un premio de la Academia de Ciencias francesa por un artículo de 1734 sobre el problema de los dos cuerpos, el cual fue desarrollado completamente por Euler en su libro *Theoria Motuum Planetarum et Cometarum*[41].

Si tenemos n cuerpos, cada uno de ellos esférico y con una masa distribuida con simetría esférica (o sea, la densidad es función del

---

[41] 1744 = *Opera*, (2), 28, 105-251.

radio), se atraerán entre sí como si sus masas estuviesen concentradas en sus centros. Sean $m_1, m_2, \ldots, m_n$ las masas y (x_i, y_i, z_i) las coordenadas (variables) de la masa i-ésima con respecto a un sistema fijo de ejes; sea r_{ij} la distancia de m_i a m_j. Entonces, las componentes según el eje x de las fuerzas que actúan sobre m_1 son

$$-\frac{k}{r_{12}^3}\, m_1 m_2\,(x_1 - x_2),\; -\frac{k}{r_{13}^3}\, m_1 m_3\,(x_1 - x_3), \ldots,$$

$$-\frac{k}{r_{1n}^3}\, m_1 m_n\,(x_1 - x_n),$$

con expresiones similares para las componentes de la fuerza según los ejes y y z, y lo mismo para cada una de las masas.

Las ecuaciones diferenciales del movimiento del cuerpo i-ésimo son entonces

$$m_i\,(d^2x_i/dt^2) = -\,km_i \sum_{j=1}^{n} m_j[(x_i - x_j)/r_{ij}^3]$$

$$m_i\,(d^2y_i/dt^2) = -\,km_i \sum_{j=1}^{n} m_j[(y_i - y_j)/r_{ij}^3] \qquad (45)$$

$$m_i\,(d^2z_i/dt^2) = -\,km_i \sum_{j=1}^{n} m_j[(z_i - z_j)/r_{ij}^3]$$

con $j \neq i$ e $i = 1, 2, \ldots, n$. Son $3n$ ecuaciones de segundo orden. Se puede elegir como origen el centro de gravedad de los n cuerpos o bien uno de ellos, por ejemplo, el sol. Hay $6n$ integrales, de las que diez se pueden calcular bastante fácilmente y son las únicas que se conocen en el caso general.

El problema de los n cuerpos, incluso para $n = 3$, de hecho, no se puede resolver exactamente; de ahí que las investigaciones sobre este problema hayan tomado dos direcciones. La primera se dirige a la búsqueda de cualesquiera teoremas de carácter general que se puedan obtener y que arrojen alguna luz sobre los movimientos, y

la segunda, a la búsqueda de soluciones aproximadas que sean útiles durante un período de tiempo subsiguiente a un instante en el que se tengan datos disponibles, lo que se conoce como método de perturbaciones.

El primer tipo de investigaciones produjo algunos teoremas sobre el movimiento del centro de gravedad de *n* cuerpos, expuestos por Newton en sus *Principia*; por ejemplo, el centro de gravedad de los *n* cuerpos se mueve con velocidad uniforme en una línea recta; las diez integrales mencionadas antes, que derivan de las denominadas leyes de conservación del movimiento, constituyen también teoremas de ese tipo y eran conocidas por Euler. Hay también algunos resultados exactos para casos especiales del problema de los tres cuerpos, debidos a uno de los maestros de la mecánica celeste, Joseph-Louis Lagrange.

Lagrange (1736-1813) era de origen francés e italiano. De niño no estaba especialmente interesado en las matemáticas, pero, estando todavía en la escuela, leyó un ensayo de Halley sobre las virtudes del cálculo de Newton y se entusiasmó por el tema. A la edad de diecinueve años se convirtió en profesor de matemáticas de la Real Escuela de Artillería de Turín, ciudad en la que nació. Pronto realizó tales contribuciones a la matemática que fue reconocido desde una edad temprana como uno de los más grandes matemáticos de la época. Aunque Lagrange trabajó en muchas ramas de la matemática —teoría de números, teoría de las ecuaciones algebraicas, cálculo infinitesimal, ecuaciones diferenciales y cálculo de variaciones— y en muchas ramas de la física, su interés principal fue la aplicación de la ley de gravitación al movimiento de los planetas. Afirmaba en 1775: «Las investigaciones aritméticas son las que me han costado más dificultades y son quizá las de menos valor.» El ídolo de Lagrange era Arquímedes.

La obra más célebre de Lagrange, su *Mécanique analytique* (1788; segunda edición, 1811-15; edición póstuma, 1853), extendió, formalizó y coronó el trabajo de Newton en mecánica. Lagrange se quejó en cierta ocasión de que Newton había sido un hombre de los más afortunados, pues había un solo universo y Newton ya había descubierto sus leyes matemáticas; sin embargo, Lagrange tuvo el honor de hacer evidente al mundo la perfección de la teoría newtoniana. Aunque la *Mécanique* es un clásico de la ciencia y también es importante para la teoría y las aplicaciones de las ecuaciones diferenciales ordinarias, Lagrange tuvo dificultades para encontrar editor.

Las soluciones particulares exactas obtenidas para el problema de tres cuerpos fueron dadas por Lagrange en un artículo premiado de 1772, *Essai sur le problème des trois corps* [42]. Una de estas soluciones establece que es posible poner tres cuerpos en movimiento de modo que sus órbitas sean elipses semejantes recorridas en el mismo tiempo y con el centro de gravedad de los tres cuerpos como un foco común. Otra solución corresponde a suponer que los tres cuerpos parten de los tres vértices de un triángulo equilátero, moviéndose entonces como si permaneciesen ligados al triángulo, que a su vez rota alrededor del centro de gravedad de los cuerpos. La tercera solución supone que los cuerpos son puestos en movimiento desde posiciones situadas sobre una línea recta; para condiciones iniciales apropiadas, continuarán sobre esa recta mientras ésta gira en un plano alrededor del centro de gravedad de los cuerpos. Para Lagrange, estos tres casos no tenían realidad física, pero en 1906 se encontró que el caso del triángulo equilátero se aplicaba al Sol, Júpiter y un asteroide llamado Aquiles.

El segundo tipo de problemas en relación con *n* cuerpos se refiere, como ya se dijo, a soluciones aproximadas, o sea, a la teoría de perturbaciones. Dos cuerpos esféricos sometidos a una atracción gravitatoria mutua se mueven a lo largo de secciones cónicas; un movimiento de este tipo se dice que es no perturbado y cualquier desviación respecto a tales movimientos, lo mismo en posición que en velocidad y sea por la causa que sea, es un movimiento perturbado. Si se trata de dos esferas pero hay resistencia por parte del medio en el cual se mueven, o si los dos cuerpos ya no son esféricos, sino, por ejemplo, esferoides achatados, o si están implicados más de dos cuerpos, entonces las órbitas ya no serán secciones cónicas. Antes de usar telescopios, las pertubaciones no eran llamativas. En el siglo XVIII, el cálculo de perturbaciones se convirtió en un problema matemático del mayor interés, al cual realizaron contribuciones Clairaut, D'Alembert, Euler, Lagrange y Laplace; el trabajo de este último en este campo fue el más destacado.

Pierre-Simon de Laplace (1749-1827) nació de padres bastante acomodados en la ciudad de Beaumont, Normandía. Pudo llegar a convertirse en sacerdote, pero se aficionó a las matemáticas en la Universidad de Caen, donde ingresó a la edad de dieciséis años; pasó

---

[42] *Hist. de l'Acad. des Sci., Paris*, 9, 1772 = *Œuvres*, 6, 229-331.

cinco años en Caen y en ese tiempo escribió un artículo sobre el cálculo de diferencias finitas. Después de acabar sus estudios, Laplace fue a París con cartas de recomendación para D'Alembert, quien no le prestó atención; Laplace escribió entonces una carta a D'Alembert exponiéndole los principios generales de la mecánica y esta vez éste le hizo caso, lo llamó y le consiguió el puesto de profesor en la École Militaire de París.

Laplace publicó prolíficamente ya de joven; una declaración hecha en la Academia de Ciencias de París poco después de su elección en 1773 señalaba que nadie tan joven había presentado tantos artículos sobre temas tan diversos y difíciles. En 1783, reemplazó a Bezout como examinador en artillería y examinó a Napoleón. Durante la Revolución fue nombrado miembro de la Comisión de Pesos y Medidas, pero fue expulsado más tarde, junto con Lavoisier y otros, por no ser un buen republicano. Laplace se retiró a Melun, una pequeña ciudad cercana a París, y allí trabajó en su célebre y popular *Exposition du système du monde* (1796). Después de la Revolución se convirtió en profesor de la École Normale, en la que también enseñaba por ese tiempo Lagrange, y formó parte de diversos comités gubernamentales. Después de eso fue ministro del Interior, miembro del Senado y canciller de dicha cámara. Aunque fue honrado por Napoleón con el título de conde, Laplace votó contra él en 1814 y se unió a Luis XVIII, quien lo nombró marqués de Laplace y par de Francia.

Durante estos años de actividad política continuó dedicándose a la ciencia. Entre 1799 y 1825 aparecieron los cinco volúmenes de su *Mécanique céleste*, obra en la que Laplace presentaba soluciones analíticas «completas» a los problemas planteados por el sistema solar; utilizó lo menos posible datos de observación; la *Mécanique céleste* incorpora descubrimientos y resultados de Newton, Clairaut, D'Alembert, Euler, Lagrange y el propio Laplace; fue una obra maestra tan completa que sus inmediatas sucesoras poco pudieron añadir; quizá su único defecto es que Laplace descuidó frecuentemente el reconocer las fuentes de sus resultados, dejando la impresión de que todos eran suyos.

En 1812 publicó su *Théorie analytique des probabilités*; la introducción a la segunda edición (1814) constituye un popular ensayo conocido como *Essai philosophique sur les probabilités*, que contiene el célebre pasaje relativo a que el futuro del mundo está completamente determinado por el pasado y que si se tuviese el conocimiento

matemático del estado del mundo en un instante dado, se podría predecir el futuro.

Laplace realizó muchos descubrimientos importantes en física matemática, algunos de los cuales consideraremos en capítulos posteriores. Realmente, estaba interesado en todo lo que contribuyese a interpretar la naturaleza; trabajó en hidrodinámica, en la propagación ondulatoria del sonido y en las mareas. En el campo de la química, es clásico su trabajo sobre el estado líquido de la materia; sus estudios sobre la tensión en la capa superficial del agua, que explica el ascenso de los líquidos en tubos capilares, y sobre las fuerzas de cohesión en los líquidos, son fundamentales. Laplace y Lavoisier diseñaron un calorímetro de hielo (1784) para medir el calor y midieron el calor específico de numerosas sustancias; el calor era todavía para ellos una clase especial de materia. La mayor parte de la vida de Laplace, sin embargo, estuvo dedicada a la mecánica celeste; murió en 1827 y sus últimas palabras se dice que fueron: «Lo que conocemos es muy poco; lo que ignoramos es inmenso.» —aunque De Morgan afirma que fueron: «El hombre sólo persigue fantasmas.»

A Laplace se le vincula con frecuencia a Lagrange, pero no se parecen ni en cualidades personales ni en su trabajo. La vanidad de Laplace le hizo abstenerse de dar el debido reconocimiento a los trabajos de aquellos a los que consideraba rivales; de hecho, utilizó muchas ideas de Lagrange sin reconocerlo, y cuando se menciona a Laplace en relación con Lagrange es siempre para elogiar las cualidades personales del segundo. Lagrange es el matemático, muy cuidadoso en sus escritos, muy claro y muy elegante. Laplace creó métodos matemáticos nuevos que posteriormente se desarrollaron en ramas de la matemática, pero nunca se preocupó de la matemática excepto en lo que le servía para estudiar la naturaleza; cuando se encontraba con un problema matemático en sus investigaciones físicas, lo resolvía casi de pasada, afirmando simplemente que «es fácil ver que...» sin molestarse en mostrar cómo había obtenido el resultado; confesaba, sin embargo, que no era fácil reconstruir su propio trabajo. Nathaniel Bowditch (1773-1838), matemático y astrónomo americano que tradujo cuatro de los cinco volúmenes de la *Mécanique céleste*, añadiendo explicaciones, dijo que cada vez que se encontraba la frase «es fácil ver que...» sabía que le esperaban horas de duro trabajo para rellenar las lagunas. En realidad, Laplace se impacientaba con las matemáticas y deseaba pasar a las aplicaciones;

la matemática pura no le interesaba y sus contribuciones a ella eran subproductos de su gran trabajo en filosofía natural; la matemática era un medio, no un fin, un útil que había que perfeccionar con el fin de resolver los problemas de la ciencia.

El trabajo de Laplace, en lo que importa al tema de este capítulo, trata de soluciones aproximadas de los problemas del movimiento planetario. La posibilidad de obtener soluciones útiles mediante un método aproximado se apoya en los siguientes factores. El sistema solar está dominado por el Sol, que contiene el 99,87 por 100 de la masa total del sistema; esto significa que las órbitas de los planetas son casi elípticas, ya que las fuerzas perturbativas entre los planetas son pequeñas; no obstante, Júpiter posee el 70 por 100 de la masa planetaria y, además, la Luna terrestre está bastante cerca de la Tierra y por ello se influyen la una a la otra. Es necesario, por todo ello, considerar perturbaciones.

El problema de los tres cuerpos, especialmente Sol, Tierra y Luna, fue muy estudiado en el siglo XVIII, en parte porque era el paso siguiente al problema de los dos cuerpos y en parte porque se necesitaba un conocimiento preciso del movimiento de la Luna para la navegación. En el caso del Sol, Tierra y Luna, se puede sacar partido del hecho de que el Sol está lejos de los otros dos cuerpos y suponer que sólo ejerce una influencia pequeña sobre el movimiento relativo de la Tierra y la Luna. En el caso del Sol y dos planetas, se considera que uno de éstos perturba el movimiento del otro alrededor del Sol; si uno de los planetas es pequeño, se puede despreciar su efecto gravitatorio sobre el otro, pero sí hay que tener en cuenta el efecto del grande sobre el pequeño. Estos casos especiales del problema de los tres cuerpos constituyen el problema restringido de los tres cuerpos.

La teoría de perturbaciones para el problema de tres cuerpos se aplicó en primer lugar al movimiento de la Luna, cosa que realizó Newton geométricamente en el tercer libro de sus *Principia*. Euler y Clairaut intentaron obtener soluciones exactas para el problema de tres cuerpos, quejándose de las dificultades y recurriendo, por supu, a métodos aproximados. Clairaut logró en esto el primer progreso real (1747) utilizando soluciones en serie de las ecuaciones diferenciales; tuvo ocasión de aplicar su resultado al movimiento del cometa Halley, que había sido observado en 1531, 1607 y 1682; se esperaba que estuviese en el perihelio de su trayectoria alrededor de la Tierra en 1759; Clairaut calculó las perturbaciones debidas a la atracción

de Júpiter y Saturno y predijo en un artículo leído en la Academia de París el 14 de noviembre de 1758 que el perihelio tendría lugar el 13 de abril de 1759, haciendo notar que el momento exacto era dudoso con un margen de un mes debido a que las masas de Júpiter y Saturno no se conocían con precisión y a ligeras perturbaciones causadas por otros planetas; el cometa alcanzó su perihelio el 13 de marzo.

Para calcular perturbaciones se desarrolló, y resultó ser el más eficaz, el denominado método de variación de los elementos o variación de los parámetros —o, todavía, de variación de las constantes de integración—. Vamos a examinarlo limitándonos a sus prinicipios matemáticos y sin tomar en consideración su trasfondo físico.

El método matemático de variación de los parámetros se remonta a los *Principia* de Newton; después de estudiar el movimiento de la Luna alrededor de la Tierra y obtener la órbita elíptica, Newton tuvo en cuenta los efectos del Sol sobre la órbita de la Luna considerando variaciones en dicha órbita. El método había sido utilizado por Jean Bernoulli en las *Acta Eruditorum* de 1697 [43] para resolver ejemplos aislados de ecuaciones no homogéneas y por Euler en 1739 al tratar la ecuación de segundo orden $y'' + k^2y = X(x)$. Fue utilizado por primera vez para estudiar perturbaciones de los movimientos planetarios por Euler en su artículo de 1748 [44], el cual trataba de las perturbaciones mutuas de Júpiter y Saturno y ganó un premio de la Academia francesa. Laplace escribió muchos artículos [45] sobre el método, método que fue desarrollado completamente por Lagrange en dos artículos [46].

Lagrange aplicó el método de variación de los parámetros para una única ecuación diferencial ordinaria a la ecuación de orden n:

$$Py + Qy' + Ry'' + \cdots + Vy^{(n)} = X,$$

donde X, P, Q, R, ... son funciones de x. Supondremos, para simplificar, que tenemos una ecuación de segundo orden.

---

[43] Página 113.

[44] *Opera*, (2), 25, 45-157.

[45] Ver, por ejemplo, *Hist. de l'Acad. des Sci., Paris*, 1772, parte 1, 651 y sigs., pub. 1775 = Œuvres, 8, 361-366, e *Hist. de l'Acad. des Sci., Paris*, 1777, 373 y sigs., pub. 1780 = *Œuvres*, 9, 357-380.

[46] *Nouv. Mém. de l'Acad. de Berlin*, 5, 1774, 201 y sigs., y 6, 1775, 190 y sigs. = *Œuvres*, 4, 5-108 y 151-251.

En el caso $X = 0$, Lagrange sabía que la solución general es

$$y = ap(x) + bq(x), \tag{46}$$

donde a y b son constantes de integración y p y q son integrales particulares de la ecuación homogénea. Ahora, dice Lagrange, consideramos a y b como funciones de x. Entonces,

$$\frac{dy}{dx} = ap' + bq' + pa' + qb'. \tag{47}$$

Lagrange hace a continuación

$$pa' + qb' = 0; \tag{48}$$

es decir, hace igual a 0 la parte de y' que resulta de la variabilidad de a y b. Resulta entonces de (47) y (48)

$$\frac{d^2y}{dx^2} = ap'' + bq'' + p'a' + q'b'. \tag{49}$$

Si la ecuación fuese de orden superior al segundo, Lagrange pondría ahora $p'a' + q'b' = 0$ y hallaría d^3y/dx^3. Como en el caso presente la ecuación es de segundo orden, conservaría en (49) todos los términos.

Sustituye ahora las expresiones de y, dy/dx y d^2y/dx^2 dadas por (46), (47) y (49) en la ecuación original; como (46) es solución de la ecuación homogénea y (48) permite eliminar alguno de los términos que resultan de la variabilidad de a y b, queda después de la sustitución

$$p'a' + q'b' = \frac{X}{R}. \tag{50}$$

Esta ecuación y la ecuación (48) constituyen un sistema algebraico en las funciones incógnitas a' y b' que se puede resolver en términos de las funciones conocidas p, q, p', q', X y R. Entonces, a y b se pueden obtener mediante una integración o al menos quedar reducidas a una cuadratura. Con estos valores de a y b, (46) da una solución de la ecuación no homogénea original; esta solución, sumada a la solución general de la ecuación homogénea, da la solución general de la ecuación no homogénea.

El método de variación de los parámetros fue tratado de una manera más general por Lagrange [47], quien demostró su aplicabilidad a muchos problemas de la física. En un artículo de 1808 lo aplicó a un sistema de tres ecuaciones de segundo orden; la técnica es, por supuesto, más complicada pero la idea es de nuevo considerar las seis constantes de integración de la solución del correspondiente sistema homogéneo como variables y determinarlas de modo que la expresión resultante satisfaga el sistema no homogéneo.

Durante el período en que estaban desarrollando el método de variación de los parámetros y en el tiempo que siguió, Lagrange y Laplace escribieron varios artículos clave sobre problemas fundamentales del sistema solar. En su obra cumbre, la *Mécanique céleste*, Laplace resumió el alcance de sus resultados:

Hemos dado, en la primera parte de este trabajo, los principios generales del equilibrio y el movimiento de cuerpos. La aplicación de estos principios a los movimientos de los cuerpos celestes nos condujo, por razonamientos geométicos [analíticos], sin ninguna hipótesis, a la ley de atracción universal, de la que son casos particulares la acción de la gravedad y el movimiento de proyectiles. Consideramos después un sistema de cuerpos sometido a esta gran ley de la naturaleza y obtuvimos, mediante un análisis apropiado, las expresiones generales de sus movimientos, de sus formas y de las oscilaciones de los fluidos que los cubren. A partir de estas expresiones hemos deducido todos los fenómenos conocidos del flujo y reflujo de las mareas, las variaciones de la gravedad en grados y en fuerza sobre la superficie de la Tierra, la precisión de los equinoccios, la libración de la Luna y la forma y rotación de los anillos de Saturno. Hemos deducido además, a partir de la misma teoría de la gravedad, las principales ecuaciones de los movimientos de los planetas, en particular, los de Júpiter y Saturno, cuyas grandes anomalías tienen un período de más de 900 años [48].

Laplace concluía que la naturaleza ordenaba la máquina celeste «para una duración eterna, sobre los mismos principios que prevalecen tan admirablemente sobre la Tierra, para la conservación de los individuos y para la perpetuación de las especies».

Según se fueron mejorando los métodos para resolver ecuaciones diferenciales y conociendo nuevos datos físicos sobre los planetas, se fueron produciendo esfuerzos a lo largo de los siglos XIX y XX

[47] *Mém. de l'Acad. des Sci., Inst. France*, 1808, 267 y sigs. = Œuvres, 6, 713-768.
[48] Prefacio al vol. 3.

para obtener mejores resultados sobre varias de las cuestiones que Laplace menciona, en particular, sobre el problema de los *n* cuerpos y la estabilidad del sistema solar.

8. Sumario

Como hemos visto, los intentos para resolver problemas físicos, que al principio involucraban solamente cuadraturas, llevaron gradualmente a darse cuenta de que se estaba creando una nueva rama de la matemática, a saber, las ecuaciones diferenciales ordinarias. A mediados del siglo XVIII las ecuaciones diferenciales se convirtieron en una disciplina independiente y su resolución en un fin en sí mismo.

La naturaleza de lo que se consideraba y buscaba como solución fue cambiando gradualmente. Al principio, los matemáticos buscaban soluciones en términos de funciones elementales, pero pronto se contentaron con dar una respuesta en términos de una cuadratura que quizá no pudiese efectuarse. Cuando fracasaron las principales vías para encontrar soluciones en términos de funciones elementales y de cuadraturas, los matemáticos se contentaron con buscar soluciones en forma de serie.

El problema de encontrar soluciones en forma cerrada no cayó en el olvido, pero en lugar de intentar resolver de ese modo las ecuaciones diferenciales concretas que surgían de los problemas físicos, los matemáticos buscaron ecuaciones diferenciales que admitiesen soluciones en términos de un número finito de funciones elementales; se encontró un gran número de ecuaciones diferenciales de este tipo. D'Alembert (1767) trabajó en este problema e incluyó las integrales elípticas entre las respuestas admisibles. Un enfoque típico de este problema, desarrollado —entre otros— por Euler (1769), consistía en empezar con una ecuación diferencial cuya integración no se podía realizar en forma cerrada y derivar de ella otra ecuación diferencial; en otro enfoque, se buscaban condiciones bajo las cuales la solución en serie pudiese contener solamente un número finito de términos.

Interesante, pero infructuoso, fue el trabajo de Marie-Jean-Antoine-Nicolas Caritat de Condorcet (1743-94) en *Du calcul intégral* intentando poner orden y método en los diversos y numerosos métodos y artificios para resolver ecuaciones diferenciales. Enumeró las operaciones de derivación, eliminación y sustitución y trató de re-

ducir todos los métodos a esas operaciones canónicas, pero su trabajo no llevó a ninguna parte. En línea con ese objetivo, Euler demostró que cuando es posible la separación de variables, existe un factor integrante, pero no recíprocamente; también demostró que la separación de variables no es aplicable a ecuaciones diferenciales de orden superior. En cuanto a las sustituciones, no halló principios generales para obtenerlas [49]; hallar sustituciones es tan difícil como resolver directamente la ecuación diferencial; sin embargo, una transformación puede reducir el orden de una ecuación diferencial y Euler sacó partido de esta idea para resolver la ecuación diferencial ordinaria lineal no homogénea de orden n, e incluso pensó que, en el caso de la ecuación homogénea, cada exp $[\int p \, dx]$, con el apropiado valor de p, daría un factor de primer orden de la ecuación diferencial ordinaria. Reducir el orden fue también el objetivo de Riccati. Se idearon otros métodos, como el de los multiplicadores de Lagrange, que se creyó que sería general pero resultó no serlo.

La búsqueda de métodos generales para integrar ecuaciones diferenciales ordinarias finalizó hacia 1775. Quedaba mucha labor por hacer en el campo de las ecuaciones diferenciales ordinarias, en particular la que se derivaba de la resolución de ecuaciones en derivadas parciales, pero no se descubrieron nuevos métodos importantes fuera de los que hemos examinado aquí hasta pasados unos cien años, cuando se introdujeron, a finales del siglo XIX, los métodos operacionales y la transformada de Laplace. En realidad, los métodos generales de resolución retrocedieron a causa de que se obtenían métodos que, de una u otra forma, se mostraban adecuados a los tipos de ecuaciones que se presentaban en las aplicaciones; no hay todavía unos principios amplios y generales para la resolución de ecuaciones diferenciales ordinarias; en su conjunto, la materia ha continuado siendo una serie de distintas técnicas para los diferentes tipos de ecuaciones.

---

[49] *Institutiones Calculi Integralis*, 1, 290.

Bibliografía

Bernoulli, Jacques: *Opera*, 2 vols., 1744, reimpreso por Birkhaüser, 1968.

Bernoulli, Jean: *Opera Omnia*, 4 vols., 1742, reimpreso por Georg Olms, 1968.

Berry, Arthur: *A Short History of Astronomy*, Dover (reimpresión), 1961, caps. 9-11.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teubner, 1898 y 1924, vol. 3 caps. 100 y 118, vol. 4, sec. 27.

Delambre, J. B. J.: *Histoire de l'astronomie moderne*, 2 vols., 1821, Johnson Reprint Corp., 1966.

Euler, Leonhard: *Opera Omnia*, Orell Füssli, serie 1, vols. 22 y 23, 1936 y 1938; serie 2, vols. 10 y 11, parte 1, 1947 y 1957.

Hofmann, J. E.: «Über Jakob Bernoullis Beiträge zür Infinitesimal-mathematik», *L'Enseignement Mathématique*, (2), 2, 61-171. Publicado separadamente por el Institut de Mathématiques, Ginebra, 1957.

Lagrange, Joseph-Louis: *Œuvres de Lagrange*, Gauthier-Villars, 1868-1873, los artículos relevantes en vols. 2, 3, 4 y 6.

—: *Mécanique analytique*, 1788, 4.ª ed., Gauthier-Villars, 1889. La cuarta edición es una reproducción inalterada de la tercera edición de 1853.

Lalande, J. de: *Traité d'astronomie*, 3 vols., 1792, Johnson Reprint Corp., 1964.

Laplace, Pierre-Simon: *Œuvres complètes*, Gauthier-Villars, 1891-1904, los artículos relevantes en los vols. 8, 11 y 13.

—: *Traité de mécanique céleste*, 5 vols., 1799-1825. También en *Œuvres complètes*, vols. 1-5, Gauthier-Villars, 1878-82. Traducción inglesa de los vols. 1-4 por Nathaniel Bowditch, 1829-39, Chelsea (reimpresión), 1966.

—: *Exposition du système du monde*, 1.ª ed., 1796, 6.ª ed. en *Œuvres complètes*, Gauthier-Villars, 1884, vol. 6.

Motucla, J. F.: *Histoire des mathématiques*, 1803, Albert Blanchard (reimpresión), 1960, vol. 3, 163-200; vol. 4, 1-125.

Todhunter, I.: *A History of the Mathematical Theories of Attraction and the figure of the Earth*, 1873, Dover (reimpresión), 1962.

Truesdell, Clifford E.: *Introduction to Leonhardi Euleri Opera Omnia, vol. X et XI Seriei Secundae*, en Euler, *Opera Omnia*, (2), 11, parte 2, Orell Füssli, 1960.

Capítulo 22
LAS ECUACIONES EN DERIVADAS PARCIALES EN EL SIGLO XVIII

> El análisis matemático es tan extenso como la propia naturaleza.
>
> JOSEPH FOURIER

1. Introducción

Lo mismo que en el caso de las ecuaciones diferenciales ordinarias, los matemáticos no crearon la disciplina de las ecuaciones en derivadas parciales de una manera consciente. Continuaron explorando los mismos problemas físicos que condujeron a aquéllas, y a medida que fueron adquiriendo una mejor comprensión de los principios físicos subyacentes a los fenómenos estudiados, formularon proposiciones matemáticas que hoy pertenecen al campo de las ecuaciones en derivadas parciales. De este modo, si los desplazamientos de una cuerda que vibra habían sido estudiados separadamente como función del tiempo y como función de la distancia de un punto de la cuerda a uno de sus extremos, el estudio de aquéllos como función de ambas variables conjuntamente y el intento de comprender todos los posibles movimientos condujeron a una ecuación en derivadas parciales. La continuación natural de este estudio, a saber, la investigación de la propagación en el aire de los sonidos generados por la cuerda, introdujo nuevas ecuaciones en derivadas parciales, abordando a continuación los matemáticos el estudio de los sonidos producidos por trompas de todas las formas, tubos de órgano, campanas, tambores y otros instrumentos.

El aire es un tipo de fluido, en el sentido en que se usa el término en física, que resulta ser compresible, mientras que los líquidos son fluidos (virtualmente) incompresibles. Las leyes del movimiento de tales fluidos y, en particular, las ondas que pueden propagarse en ambos tipos de fluidos dieron lugar a un extenso campo de investigación que hoy constituye la disciplina de la hidrodinámica; también en este campo surgieron ecuaciones en derivadas parciales.

A lo largo del siglo XVIII, los matemáticos continuaron trabajando sobre el problema de la atracción gravitatoria ejercida por cuerpos de diferentes formas, especialmente el elipsoide. Si bien es éste, básicamente, un problema de integración triple, Laplace lo convirtió en un problema de ecuaciones en derivadas parciales de una manera que examinaremos en breve.

2. La ecuación de ondas

Aunque aparecieron ecuaciones en derivadas parciales concretas ya en 1734 en la obra de Euler [1] y en 1743 en el *Traité de dynamique* de D'Alembert, no se hizo con ellas nada que merezca la pena reseñar. El primer éxito real con las ecuaciones en derivadas parciales llegó como fruto de renovados ataques al problema de la cuerda vibrante, tipificada por una cuerda de violín. La aproximación de que las vibraciones sean pequeñas se impuso para hacer tratable la ecuación en derivadas parciales. Jean Le Rond d'Alembert (1717-83), en sus artículos de 1746 [2] titulados «Investigaciones sobre la curva que forma una cuerda tensa que se hace vibrar», afirma que se propone demostrar que existen infinitas curvas además de la curva seno que son modos de vibración.

Recordemos del capítulo precedente que en las primeras aproximaciones al problema de la cuerda vibrante, ésta se contemplaba como un «collar de cuentas», es decir, la cuerda se consideraba compuesta de una cantidad discreta de *n* pesos iguales e igualmente espaciados, unidos por trozos de hilo elástico, flexible y sin peso. Para tratar la cuerda continua, se hacía tender a infinito el número de pesos mientras decrecía el tamaño y la masa de cada uno de ellos, de manera que la masa total de la creciente cantidad de «cuentas»

---

[1] *Comm. Acad. Sci. Petrop.*, 7, 1734/35, 184-200, pub. 1740 = *Opera*, (1), 22, 57-75.

[2] *Hist. de l'Acad de Berlin*, 3, 1747, 214-219 y 220-249, pub. 1749.

individuales se aproximase a la masa de la cuerda continua; había dificultades matemáticas en ese paso al límite, pero esas sutilezas se pasaban por alto.

El caso de una cantidad discreta de masas había sido tratado por Jean Bernoulli en 1727 (cap. 21, sec. 4); si la cuerda tiene longitud l y ocupa $0 \leq x \leq l$, y si x_k es la abcisa de la masa k-ésima, $k = 1, 2, \ldots, n$ (la masa n-ésima, situada en $x = l$, no se mueve), entonces

$$x_k = k\frac{l}{n}, \qquad k = 1, 2, \ldots, n.$$

Analizando la fuerza que actúa sobre la masa k-ésima, Bernoulli había probado que si y_k es el desplazamiento de dicha masa, entonces

$$\frac{d^2y_k}{dt^2} = \left(\frac{na}{l}\right)^2 (y_{k+1} - 2y_k + y_{k-1}), \qquad k = 1, 2, \ldots, n-1$$

donde $a^2 = lT/M$, T es la tensión de la cuerda (que se supone constante mientras la cuerda oscila), y M es la masa total. D'Alembert reemplazó y_k por $y(t, x)$ y l/n por Δx, con lo que

$$\frac{\partial^2 y(t, x)}{\partial t^2} = a^2\left[\frac{y(t, x+\Delta x) - 2y(t, x) + y(t, x-\Delta x)}{(\Delta x)^2}\right],$$

observando a continuación que cuando n se hace infinito, o sea, cuando Δx tiende a 0, la expresión entre corchetes converge a $\partial^2 y/\partial x^2$, de donde

$$\frac{\partial^2 y(t, x)}{\partial t^2} = a^2\frac{\partial^2 y(t, x)}{\partial x^2}, \tag{1}$$

siendo ahora $a^2 = T/\sigma$ y σ la masa por unidad de longitud. Aparece así por primera vez la que hoy se conoce como ecuación de ondas en una dimensión espacial.

Como la cuerda está fijada en los extremos $x = 0$ y $x = l$, la solución ha de satisfacer las condiciones de contorno

$$y(t, 0) = 0, \qquad y(t, l) = 0. \tag{2}$$

En el instante $t = 0$ se fuerza a que la cuerda adopte un perfil $y = f(x)$ y entonces se suelta, lo que significa que cada partícula comienza

con una velocidad inicial nula; estas condiciones iniciales se expresan matemáticamente en la forma

$$y(0, x) = f(x) \qquad \left.\frac{\partial y(t, x)}{\partial t}\right|_{t=0} = 0 \tag{3}$$

y también han de ser satisfechas por la solución.

Este problema fue resuelto por D'Alembert de una manera tan ingeniosa que se reproduce con frecuencia en los textos modernos. No nos detendremos en todos los detalles; demostró en primer lugar que

$$y(t, x) = \frac{1}{2}\,\phi(at + x) + \frac{1}{2}\,\psi(at - x), \tag{4}$$

donde ϕ y ψ son dos funciones que todavía hay que determinar; con ello, D'Alembert dedujo que *toda* solución de la ecuación (1) es la suma de una función de $(at + x)$ y una función de $(at - x)$; el recíproco es fácil de probar por sustitución directa de (4) en (1). La solución ha de satisfacer las condiciones iniciales y de contorno; imponiendo la condición $y(t, 0) = 0$ a (4) se tiene, para todo t

$$\frac{1}{2}\,\phi(at) + \frac{1}{2}\,\psi(at) = 0. \tag{5}$$

Puesto que cualquiera que sea x se tiene que $ax + t = at'$ para cierto valor de t', podemos afirmar que

$$\phi(x + at) = -\ \psi(x + at), \tag{6}$$

cualesquiera que sean x y t. En consecuencia, la condición $y(t, l) = 0$ se convierte, a la vista de (4), en

$$\frac{1}{2}\,\phi(at + l) = \frac{1}{2}\,\phi(at - l); \tag{7}$$

y como esta última es una identidad en t, se deduce que ϕ ha de ser periódica en $at + x$ con período $2l$.

La condición

$$\left.\frac{\partial y(t, x)}{\partial t}\right|_{t=0} = 0 \tag{8}$$

implica, por (4) y el hecho de que $\phi = -\psi$,

$$\phi'(x) = \phi'(-x), \tag{9}$$

lo que, integrando, da

$$\phi(x) = -\phi(-x), \tag{10}$$

y por lo tanto ϕ es una función impar de x. Si utilizamos ahora en (4) el que $\phi = -\psi$, formamos $y(0, x)$ y aplicamos (10), se tendrá que

$$y(0, x) = \phi(x), \tag{11}$$

y como la condición inicial es $y(0, x) = f(x)$, resultará

$$\phi(x) = f(x) \text{ para } 0 \leq x \leq l. \tag{12}$$

Resumiendo,

$$y(t, x) = \frac{1}{2}\phi(at + x) - \frac{1}{2}\phi(at - x), \tag{13}$$

donde ϕ ha de satisfacer las anteriores condiciones de periodicidad e imparidad; además, si el estado inicial es $y(0, x) = f(x)$, se ha de cumplir (12) entre 0 y l. Habrá, así, una única solución para cada $f(x)$ dada. D'Alembert consideraba las funciones como expresiones analíticas formadas mediante los procesos del álgebra y el cálculo infinitesimal, con lo que si dos de tales funciones coincidían en un intervalo de valores de la x, habrían de coincidir para todo valor de la x; puesto que $\phi(\ x) = f(x)$ en $0 \leq x \leq l$ y ϕ ha de ser impar y periódica, lo mismo ha de cumplir $f(x)$. Finalmente, como $y(t, x)$ ha de satisfacer la ecuación diferencial, ha de tener derivadas segundas; pero $y(0, x) = f(x)$, con lo que $f(x)$ habrá de poseer derivadas segundas.

Pocos meses después de ver los artículos de 1746 de D'Alembert, Euler escribió por su parte el artículo «Sobre la oscilación de cuerdas», que fue presentado el 16 de mayo de 1748 [3]. Aunque en el método de solución siguió a D'Alembert, Euler tenía por ese tiempo

[3] *Nova Acta Erud.*, 1749, 512-527 = *Opera*, (2), 10, 50-62; también en francés por Euler, *Hist. de l'Acad. de Berlin*, 4, 1748, 69-85 = *Opera*, (2), 10, 63-77.

una idea completamente distinta en cuanto a qué funciones se podían admitir como curvas iniciales y, en consecuencia, como soluciones de una ecuación en derivadas parciales. Antes incluso del debate sobre el problema de la cuerda vibrante, en un trabajo de 1734, de hecho Euler aceptaba funciones formadas a partir de trozos de curvas conocidas e incluso las obtenidas dibujando curvas a pulso. Así, la curva (fig. 22.1) formada por un arco de parábola en el intervalo (a, c) y por un arco de una curva de tercer grado en el intervalo (c, d) constituía una curva o función en su concepto. Euler denominaba discontinuas a tales curvas, aunque en terminología moderna son continuas con derivada discontinua. En su texto, la *Introductio* de 1748, se adhirió a la noción que era corriente en el siglo XVIII, a saber, que una función ha de estar dada por una única expresión analítica. Pero fueron, al parecer, argumentos de naturaleza física en relación con el problema de la cuerda vibrante los que le impulsaron a anteponer su nuevo concepto de función, aceptando cualquier función definida por una fórmula $\phi(x)$ en $-l \leq x \leq l$, y tomando $\phi(x + 2l) = \phi(x)$ como definición de la curva fuera de $(-l, l)$. En un artículo posterior [4], Euler va más allá, afirmando que

$$y = \phi(ct + x) + \psi(ct - x), \tag{14}$$

con ϕ y ψ arbitrarias, es solución de

$$\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}, \tag{15}$$

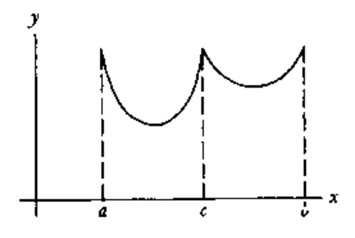


FIGURA 22.1

[4] *Hist. de l'Acad de Berlin*, 9, 1753, 196-222, pub. 1755 = *Opera*, (2), 10, 232-254.

lo que se sigue por sustitución en la ecuación diferencial; pero la curva inicial es igualmente admisible si se expresa por una ecuación que si se traza de una manera no expresable por una ecuación; de la curva inicial, sólo es relevante la parte que está en $0 \leq x \leq l$, y no ha de tomarse en consideración la continuación de esa parte. Resulta así que las distintas partes de dicha curva no están conectadas entre sí mediante una ley de continuidad (una expresión analítica única), sino mediante la descripción de más arriba; por esta razón, puede ser imposible incluir la curva completa en una ecuación, excepto cuando por casualidad la curva está dada por una función sinusoidal.

En 1755, Euler dio como nueva definición de función la siguiente: «Si unas cantidades dependen de otras de tal modo que sufren una variación cuando estas últimas varían, entonces se dice que las primeras son funciones de las segundas.» Y en otro artículo [5] afirma que las partes de una función «discontinua» no se pertenecen unas a otras y no están determinadas por una única ecuación válida para la función en toda su extensión. Además, dado el perfil inicial en $0 \leq x \leq l$, se repite en orden inverso en $-l \leq x \leq 0$ (para tener una función impar), imaginando después una repetición continua de la curva resultante en cada intervalo de longitud $2l$, hasta el infinito. Entonces, si se utiliza esa curva $[y = f(x)]$ para representar la función inicial, la ordenada en el instante t correspondiente a la abscisa x de la cuerda que oscila, estará dada por (cf. (13) y (12))

$$y = \frac{1}{2} f(x + ct) + \frac{1}{2} f(x - ct). \tag{16}$$

En su fundamental artículo de 1749, Euler señala que todos los posibles movimientos de la cuerda vibrante son periódicos respecto al tiempo, cualquiera que sea el perfil de la cuerda; el período es (normalmente) el período de lo que hoy denominamos modo fundamental. También se apercibió de que pueden presentarse modos individuales cuyos períodos son un medio, un tercio, y así sucesivamente, del período fundamental, dando para tales soluciones especiales la expresión

$$y(t, x) = \sum A_n \operatorname{sen} \frac{n\pi x}{l} \cos \frac{n\pi ct}{l} \tag{17}$$

[5] *Novi Comm. Acad. Sci. Petrop.*, 11, 1765, 67-102, pub. 1767 = *Opera*, (1), 23, 74-91.

cuando el perfil inicial es

$$y(0, x) = \sum A_n \operatorname{sen} \frac{n\pi x}{l}, \tag{18}$$

aunque no aclara si el signo de sumación se refiere a un número finito o infinito de términos. Euler tiene, no obstante, la idea de la superposición de modos. Así pues, el principal punto de desacuerdo de Euler con D'Alembert es que el primero admitiría toda clase de curvas iniciales, y por tanto soluciones no analíticas, mientras que D'Alembert aceptaba únicamente soluciones y curvas iniciales analíticas.

Al introducir sus funciones «discontinuas», Euler comprendió que había dado un gran paso hacia adelante; el 20 de diciembre de 1720 escribió a D'Alembert que «el considerar tal tipo de funciones, no sujetas a ninguna ley de continuidad [analiticidad], nos abre todo un nuevo campo de análisis» [6].

El problema de la cuerda vibrante fue resuelto de una manera completamente diferente por Daniel Bernoulli; su trabajo suscitó otro motivo para la controversia relativa a las soluciones admisibles. Daniel Bernoulli (1700-82), hijo de Jean Bernoulli, fue profesor de matemáticas en San Petersburgo de 1725 a 1733 y después, sucesivamente, profesor de medicina, metafísica y filosofía natural en Basilea; realizó su obra más importante en hidrodinámica y elasticidad, ganando, en el primero de estos campos, un premio por un artículo sobre el flujo de las mareas; también consideró la aplicación de la teoría sobre el movimiento de líquidos al flujo de la sangre humana en los vasos sanguíneos. Fue un hábil experimentador y descubrió antes de 1760, mediante trabajo experimental, la ley de atracción de cargas eléctricas estáticas, ley que habitualmente se atribuye a Charles Coulomb. La *Hydrodinamica* de Bernoulli (1738), que contiene estudios que aparecieron en diversos artículos, es el primer texto importante en su campo; incluye un capítulo sobre la teoría mecánica del calor (en tanto que opuesta a la consideración del calor como una sustancia) y da muchos resultados sobre la teoría de gases.

En su artículo de 1732-33 citado en el capítulo precedente, Daniel Bernoulli había afirmado expresamente que la cuerda vibrante

[6] *Opera*, (2), 11, sec. 1, 2.

podía tener modos superiores de oscilación; en un artículo [7] posterior sobre la oscilación de una cuerda vertical flexible cargada con pesos, hizo la siguiente observación:

Análogamente, una cuerda musical tensa puede producir sus vibraciones isocronas de muchas maneras, incluso, de acuerdo con la teoría, de infinitas maneras..., y además en cada momento emite una nota superior o inferior. El primer y más natural modo tiene lugar cuando la cuerda produce en sus oscilaciones un sólo arco; se produce entonces la oscilación más lenta y se emite el tono más bajo de entre todos los posibles, fundamental respecto al resto. El siguiente modo exige que la cuerda produzca dos arcos a lados opuestos [de la posición de equilibrio de la cuerda] siendo entonces la oscilación el doble de rápida y emitiéndose ahora la octava del sonido fundamental.

Describe después los modos superiores; no proporciona los argumentos matemáticos pertinentes, pero parece evidente que disponía de ellos.

En un artículo sobre las vibraciones de una barra y los sonidos que ésta emite [8], Bernoulli no sólo da los distintos modos en que la barra puede vibrar sino que afirma claramente que ambos sonidos (el fundamental y un armónico superior) pueden existir a la vez. Es ésta la primera afirmación de la coexistencia de oscilaciones armónicas pequeñas; Bernoulli la basó en su comprensión física de cómo pueden comportarse la barra y los sonidos, pero no dio ningún argumento matemático para justificar que la suma de dos modos es una solución.

Cuando leyó el primer artículo de D'Alembert de 1746 y el de Euler de 1749 sobre la cuerda vibrante, se apresuró a publicar las ideas que había obtenido durante muchos años [9]. Después de permitirse sarcasmos sobre el carácter abstracto de los trabajos de D'Alembert y Euler, reitera que pueden existir simultáneamente muchos modos de oscilación en la cuerda vibrante (ésta responde entonces a la suma o superposición de todos los modos) y pretende que esto es todo lo que Euler y D'Alembert habían demostrado. Plantea después una cuestión fundamental; insiste en que *todas* las posibles curvas iniciales se pueden representar en la forma.

---

[7] *Comm. Acad. Sci. Petrop.*, 12, 1740, 97-108, pub. 1750.
[8] *Comm. Acad. Sci. Petrop.*, 13, 1741/43, 167-196, pub. 1751.
[9] *Hist. de l'Acad. de Berlin*, 9, 1753, 147-172 y 173-195, pub. 1755.

$$f(x) = \sum_{n=1}^{\infty} a_n \operatorname{sen} \frac{n\pi x}{l}, \tag{19}$$

porque existen suficientes constantes a_n como para que la serie se ajuste a cualquier curva. En consecuencia, afirma, *todos* los correspondientes movimientos vendrán dados por

$$y(t, x) = \sum_{n=1}^{\infty} a_n \operatorname{sen} \frac{n\pi x}{l} \cos \frac{n\pi ct}{l}. \tag{20}$$

Así pues, cualquier movimiento, correspondiente a una curva inicial, no es más que una suma de modos periódicos sinusoidales, y la combinación tiene la frecuencia del modo fundamental. Sin embargo, Bernoulli no dio argumentos matemáticos para apoyar sus afirmaciones; se apoyó en argumentos físicos. En este artículo de 1753 afirmaba:

Mi conclusión es que todos los cuerpos sonoros incluyen una infinidad de sonidos con una correspondiente infinidad de vibraciones regulares... Pero no es de esta multitud de sonidos de la que los señores D'Alembert y Euler dicen hablar... Cada clase [cada modo fundamental generado por una curva inicial] se multiplica un número infinito de veces para ajustar a cada intervalo un número infinito de curvas, de modo que cada punto comienza y termina en el mismo instante estas vibraciones mientras que, de acuerdo con la teoría del señor Taylor, cada intervalo entre dos nodos adoptará la forma de la asociada de la cicloide [la función seno] muy alargada.

Señalemos además que la cuerda *AB* no puede vibrar sólo conforme a la primera figura [modo fundamental] o la segunda [segundo armónico] o la tercera, y así hasta el infinito, sino que podrá hacerlo según una combinación de esas vibraciones entre todas las combinaciones posibles, y que todas las nuevas curvas dadas por D'Alembert y Euler son únicamente combinaciones de las vibraciones de Taylor.

En esta última observación, Bernoulli atribuye a Taylor algo que éste nunca demostró conocer. Aparte de esto, las afirmaciones de Bernoulli son enormemente importantes.

Euler puso inmediatamente objeciones a la última afirmación de Bernoulli. De hecho, el artículo de Euler de 1753 presentado a la Academia de Berlín (ya mencionado más arriba) era en parte una réplica a los dos artículos de Bernoulli. Euler resalta la importancia

de la ecuación de ondas como punto de partida para el estudio de la cuerda vibrante; elogia cómo Bernoulli vio que pueden existir simultáneamente muchos modos de vibración y, por ende, que la cuerda puede emitir en un movimiento muchos armónicos, pero niega, lo mismo que D'Alembert, que todos los posibles movimientos puedan representarse por (20). Admite que una curva inicial tal como

$$f(x) = \frac{c \operatorname{sen}(ax/l)}{1 - a \cos(ax/l)}, \quad |a| < 1, \tag{21}$$

puede expresarse por una serie como (19); Bernoulli estaría en lo cierto si toda función se pudiera expresar mediante una serie trigonométrica, pero Euler considera que esto es imposible. Una suma de funciones seno es, afirma Euler, una función periódica impar, pero en su solución (ver [16]),

$$y(t, x) = \frac{1}{2} f(x + ct) + \frac{1}{2} f(x - ct), \tag{22}$$

f es arbitraria (discontinua, en el sentido de Euler), con lo que no puede expresarse como una suma de funciones seno. De hecho, afirma, podría ser una combinación de arcos extendidos sobre el dominio infinito de las x y ser impar y periódica, pero aun así, a causa de ser discontinua (en el sentido de Euler), no podría expresarse como suma de curvas seno. Su propia solución, afirma, no tiene ningún tipo de limitación y, en efecto, la curva inicial no necesita poder expresarse mediante una ecuación (una única expresión analítica).

Euler se refirió también, en este caso correctamente, a la serie de Maclaurin afirmando que ésta no podía representar cualquier función arbitraria; en consecuencia, tampoco podía hacerlo una serie de senos. Todo lo más, concedía que las series trignométricas de Bernoulli representaban soluciones particulares y, en efecto, el propio Euler había obtenido tales soluciones en su artículo de 1749 (ver [17] y [18]).

D'Alembert, en su artículo «Fondamental» del volumen 7 (1757) de la *Encyclopédie*, atacó también a Bernoulli. No creía que todas las funciones periódicas e impares se pudiesen representar mediante una serie como (19), pues la serie es derivable dos veces mientras que no toda función de aquel tipo tiene por qué serlo. No obstante,

incluso cuando la curva inicial es suficientemente derivable y D'Alembert requería en su artículo de 1746 que fuese dos veces derivable— no tiene por qué ser representable en la forma de Bernoulli. D'Alembert puso también reparos, sobre la misma base, a las curvas discontinuas de Euler. De hecho, el requisito de D'Alembert de que la curva inicial $y = f(x)$ sea dos veces derivable era correcto, pues una solución obtenida a partir de una $f(x)$ que no tenga derivada segunda en ciertos valores de x ha de satisfacer condiciones especiales en tales puntos singulares.

Bernoulli no se rectractó de su postura. En una carta de 1758 [10] repite que tiene en los a_n una cantidad infinita de coeficientes a su disposición que, elegidos apropiadamente, le permiten hacer coincidir la serie de (19) con cualquier función en una cantidad infinita de puntos. En cualquier caso, Bernoulli insistió en que (20) era la solución más general. La discusión entre D'Alembert, Euler y Bernoulli continuó por una década sin que se alcanzase un acuerdo. La esencia del problema era la amplitud de la clase de funciones que se podían representar mediante una serie de senos, o, más generalmente, una serie de Fourier.

En 1759, Lagrange, entonces joven y desconocido, se unió a la controversia. En su artículo, que trataba de la naturaleza y propagación del sonido [11], daba algunos resultados sobre este tema y después aplicaba su método a la cuerda vibrante. Procedió como si estuviese abordando un problema nuevo, aunque repitió mucho de lo que habían hecho antes Euler y Daniel Bernoulli. Lagrange comenzó también con una cuerda cargada con un número finito de masas iguales e igualmente espaciadas, pasando después al límite cuando ese número se hace infinito. Aunque criticó el método de Euler por restringir los resultados a curvas continuas (analíticas), Lagrange afirmó que demostraría que la conclusión de Euler —que cualquier curva inicial puede servir—, era correcta. Pasemos inmediatamente a la conclusión de Lagrange para la cuerda continua; había obtenido

$$y(x, t) = \frac{2}{l} \sum_{r=1}^{\infty} \operatorname{sen} \frac{r\pi x}{l} \sum_{q=1}^{\infty} \operatorname{sen} \frac{r\pi x}{l}\, dx$$

[10] *Jour. des Sçavans*, marzo 1758, 157-166.
[11] *Misc. Taur.*, 1_3, 1759, i-x, 1-112 = *Œuvres*, 1, 39-148.

$$\left(Y_q \cos \frac{c\pi rt}{l} + \frac{l}{r\pi c} V_q \operatorname{sen} \frac{c\pi rt}{l}\right), \tag{23}$$

donde Y_q y V_q son el desplazamiento inicial y la velocidad inicial de la masa q-ésima. Después, reemplazó Y_q y V_q por $Y(x)$ y $V(x)$, respectivamente. Lagrange consideraba las cantidades

$$\sum_{q=1}^{\infty} \operatorname{sen} \frac{r\pi x}{l} Y(x)\, dx \qquad y \qquad \sum_{q=1}^{\infty} sen \frac{r\pi x}{l} V(x)\, dx \tag{24}$$

como integrales y sacó la operación integración fuera del sumatorio $\Sigma_{r=1}^{\infty}$, resultando de todo ello

$$y(t, t) = \left(\frac{2}{l}\int_0^l Y(x) \sum_{r=1}^{\infty} \operatorname{sen} \frac{r\pi x}{l}\, dx\right) \operatorname{sen} \frac{r\pi x}{l} \cos \frac{r\pi ct}{l}$$

$$+ \left(\frac{2}{\pi c}\int_0^l V(x) \sum_{r=1}^{\infty} \frac{1}{r} \operatorname{sen} \frac{r\pi x}{l}\, dx\right) \operatorname{sen} \frac{r\pi x}{l} \operatorname{sen} \frac{r\pi ct}{l}. \tag{25}$$

El intercambio entre integral y sumatorio no sólo introdujo series divergentes, sino que echó a perder toda posiblidad de que Lagrange pudiese haber identificado

$$\int_0^l Y(x) \operatorname{sen} \frac{r\pi x}{l}\, dx \tag{26}$$

como un coeficiente de Fourier. Después de otros largos, difíciles y dudosos pasos, Lagrange obtuvo el resultado de Euler y D'Alembert

$$y = \phi(ct + x) + \psi(ct - x), \tag{27}$$

y concluía que la anterior deducción colocaba la teoría de este gran geómetra [Euler]

más allá de toda duda y establecida sobre principios directos y claros que no reposan en aboluto sobre la ley de continuidad [analiticidad] que D'A-

lembert requiere; así es, además, cómo puede ocurrir que la misma fórmula que sirve para respaldar y probar la teoría de Bernoulli sobre la mezcla de vibraciones isocronas cuando el número de cuerpos es... finito nos muestra su insuficiencia... cuando el número de esos cuerpos se hace infinito. En efecto, el cambio que sufre esta fórmula al pasar de un caso a otro es tal que los movimientos absolutos simples que componen los movimientos absolutos del sistema completo se anulan entre sí en su mayor parte, y los que quedan están tan desfigurados y alterados que resultan completamente irreconocibles. Verdaderamente, es una contrariedad que una teoría tan ingeniosa... resulte falsa en el caso principal, con el cual están relacionados todos los movimientos recíprocos pequeños que se presentan en la naturaleza.

Todo esto es un absurdo casi completo.

El principal argumento de Lagrange para sostener que su solución no exige restricciones sobre la curva inicial $Y(x)$ y la velocidad inicial $V(x)$ es que no aplica a éstas la operación de derivación. Pero si intentásemos rigorizar lo que él hizo sí habría que imponer restricciones.

Euler y D'Alembert criticaron el trabajo de Lagrange, pero en realidad apuntaron a detalles de sus «prodigiosos cálculos», como dijo Euler, y no señalaron sus principales fallos. Lagrange intentó responder a las críticas, siendo las réplicas y refutaciones de ambos campos demasiado extensas para relatarlas aquí, aunque muchas de ellas son reveladoras del pensamiento de la época. Por ejemplo, Lagrange reemplazó sen π/m para $m=\infty$ por π/m y sen $v\pi/2m$ por $v\pi/2m$ *para* $m=\infty$. D'Alembert aceptaba lo primero pero no lo segundo, porque los valores de v implicados eran comparables a los de m. La objeción de que una serie de la forma

$$\cos x + \cos 2x + \cos 3x + \cdots$$

podía ser divergente fue también suscitada por D'Alembert, replicando Lagrange con el argumento, común en aquel tiempo, de que el valor de la serie es el valor de la función de la que la serie proviene.

Aunque Euler criticó algunos detalles matemáticos, su juicio de conjunto sobre el artículo de Lagrange, comunicado en una carta del 23 de octubre de 1759 [12], fue elogiar la destreza matemática de Lagrange y afirmar que ponía toda la discusión más allá de cualquier objeción, y que todo el mundo debería ya aceptar el uso de funcio-

[12] Lagrange, *Œuvres*, 14, 164-170.

nes irregulares y discontinuas (en el sentido de Euler) en esta clase de problemas.

El 2 de octubre de 1759, Euler escribió a Lagrange: «Estoy encantado de saber que está usted de acuerdo con mi solución... que D'Alembert ha intentado socavar con diversos reparos por la sola razón de que no la obtuvo él mismo. Ha amenazado con publicar una refutación de peso; no sé si realmente lo hará; él piensa que podrá engañar a los semicultos con su elocuencia. Dudo que sea serio, a no ser que esté, quizá, completamente cegado por al egolatría.» [13]

En 1760-61, Lagrange, intentando responder a las críticas que D'Alembert y Bernoulli habían comunicado por carta, dio una solución diferente al problema de la cuerda vibrante [14]. Esta vez, Lagrange arranca directamente de la ecuación de ondas (con $c = 1$) y, multiplicando por una función incógnita junto con otros pasos adicionales, reduce la ecuación en derivadas parciales a la resolución de dos ecuaciones diferenciales ordinarias. Mediante nuevos pasos, no todos correctos, Lagrange obtiene la solución

$$y(t, x) = \frac{1}{2} \quad f(x+t) + \quad \frac{1}{2} \quad f(x-t) - \quad \frac{1}{2}\int_0^{x+t} g\, dx + \frac{1}{2}\int_0^{x-t} g\, dx,$$

en donde $f(x) = y(0,x)$ y $g(x) = \partial y/\partial t$ en $t = 0$ son los datos iniciales. Este resultado coincide con el de D'Alembert, según muestra Lagrange. Pero después, sin referirse a su propio trabajo, Lagrange trata de convencer a sus lectores de que no había usado ninguna ley de continuidad (analiticidad) para la curva inicial; es cierto que no utilizó ninguna operación directa de derivación sobre la función inicial, pero para justificar rigurosamente sus procedimientos de paso al límite no se puede evitar, tampoco en este artículo, el hacer hipótesis relativas a la continuidad y la derivabilidad de las funciones iniciales.

El debate continuó con pleno vigor durante todos los años sesenta y setenta del siglo. Incluso Laplace entró en la refriega en

[13] *Œuvres*, 14, 162-164.

[14] *Misc. Taur.*, 2_2, 1760/61, 11-172, pub. 1762 = *Œuvres*, 1, 151-316.

1779 [15], poniéndose al lado de D'Alembert; éste prosiguió en una serie de folletos, titulados *Opuscules*, que comenzaron a aparecer en 1768. Polemizó contra Euler porque éste admitía curvas iniciales demasiado generales y contra Bernoulli porque su solución (la de D'Alembert) no se podía representar como suma de curvas seno, con lo que las soluciones de Bernoulli no eran suficientemente generales. La idea de que una serie de funciones trigonométricas $\Sigma_{n=1}^{\infty} a_n$ sen nx se podía ajustar a cualquier curva inicial gracias a que hay una infinidad de a_n a determinar (según había sostenido Daniel Bernoulli) fue rechazada por Euler como imposible de llevar a cabo. También suscitó la cuestión de cómo una serie trigonométrica podría representar la curva inicial cuando sólo se perturba inicialmente una parte de la cuerda. Euler, D'Alembert y Lagrange continuaron negando hasta el fin que una serie trigonométrica pudiese representar cualquier función analítica, por no decir nada de funciones más arbitrarias.

Muchos de los argumentos presentados por cada uno de ellos eran muy incorrectos y los resultados obtenidos en el siglo XVIII no fueron definitivos. Una de las principales cuestiones, la representabilidad de una función arbitraria mediante una serie trigonométrica, no se aclaró hasta que Fourier se ocupó de ella. Euler, D'Alembert y Lagrange estuvieron a punto de descubrir la importancia de las series de Fourier, pero no llegaron a apreciar lo que tenían delante. Tanto ellos tres como Bernoulli tenían razón, juzgando por los conocimimientos de la época, en sus principales aseveraciones. D'Alembert, siguiendo una tradición establecida desde los tiempos de Leibniz, insistía en que todas las funciones habían de ser analíticas, de tal manera que un problema que no se pudiera resolver en tales términos era irresoluble. Tuvo razón en el argumento de que $y(t, x)$ ha de ser periódica en x, pero no vio que, dada una función arbitraria en, por ejemplo, $0 \leq x \leq l$, dicha función se puede repetir en cada intervalo $[nl, (n+1)l]$, con n entero, y resultar así periódica. Por supuesto, puede que tal función periódica no sea representable mediante una fórmula cerrada. Euler y Lagrange tenían razones, al menos en su tiempo, para creer que no toda función «discontinua» se puede representar por una serie de Fourier, y también tenían razón al creer (aunque no disponían de una demostración) que la curva inicial podía ser muy general; no necesita ser analítica ni tam-

[15] *Mém. de l'Acad des Sci., Paris*, 1779, 207-309, pub. 1782 = *Œuvres*, 10, 1-89.

poco ser periódica. Bernoulli fue quien mantuvo la postura correcta con argumentos físicos, pero no la pudo justificar matemáticamente.

Uno de los aspectos ciertamente curiosos del debate sobre la representación de funciones mediante series trigonométricas es que todos los implicados sabían que se puede representar (en un intervalo) funciones no periódicas por tales series. Una consulta al capítulo 20 (sec. 5) mostrará que Clairaut, Euler, Daniel Bernoulli y otros habían obtenido de hecho tales representaciones; muchos de sus artículos traían también las fórmulas para los coeficientes de las series trigonométricas. Estos trabajos estaban impresos prácticamente todos en 1759, el año en que Lagrange presentó su fundamental artículo sobre la cuerda vibrante. Podía éste, por tanto, haber inferido de dichos trabajos que toda función admite un desarrollo en serie trigonométrica y haber tomado de ellos las fórmulas para los coeficientes, pero no lo hizo. No fue hasta 1773, cuando el ardor de la controversia había pasado, cuando Daniel Bernoulli hizo observar que la suma de una serie trigonométrica puede representar distintas expresiones algebraicas en intervalos distintos. ¿Por qué todos estos resultados no ejercieron ninguna influencia sobre la controversia relativa a la cuerda vibrante? Se puede explicar de diversas maneras. Muchos de los resultados sobre la representación en serie trigonométrica de funciones bastante generales aparecieron en artículos sobre astronomía y pudo ocurrir que Daniel Bernoulli no los hubiese leído, no pudiendo así recurrir a ellos para defender su postura. Euler y D'Alembert, que tuvieron que conocer el trabajo de Clairaut de 1757 (cap. 20, sec. 5), no estaban probablemente inclinados a estudiarlo ya que refutaba sus propios argumentos y, además, ese trabajo astronómico de Clairaut fue pronto superado y olvidado. Por otra parte, aunque Euler utilizó series trigonométricas, como en su trabajo sobre teoría de la interpolación, para representar expresiones polinomiales, no aceptó el resultado general de que funciones muy generales se pudiesen representar así; la existencia de tales representaciones en serie, cuando las utilizaba, se aseguraba por otros medios.

Otra cuestión, la de cómo una ecuación en derivadas parciales con coeficientes analíticos (por ejemplo, constantes) podía tener una solución no analítica, no llegó realmente a clarificarse. En el caso de ecuaciones diferenciales ordinarias, si los coeficientes son analíticos, las soluciones también lo han de ser, pero esto no es cierto para ecuaciones en derivadas parciales. Aunque Euler tenía razón al afir-

mar que son admisibles soluciones angulosas (e insistió en ello), la determinación de las singularidades que son admisibles en la solución de una ecuación en derivadas parciales estaba todavía lejos.

3. Extensiones de la ecuación de ondas

Mientras continuaba la controversia sobre la cuerda vibrante, el interés en los instrumentos musicales propició nuevos trabajos no sólo sobre las vibraciones de estructuras físicas, sino también sobre cuestiones hidrodinámicas concernientes a la propagación del sonido en el aire. Ello implica, en lo matemático, extensiones de la ecuación de ondas.

En 1762, Euler abordó el problema de la cuerda vibrante con grosor variable. Había sido estimulado a ello por una de las principales cuestiones de la estética musical: Jean-Philippe Rameau (1683-1764) había explicado en 1726 que la consonancia de un sonido musical es debida al hecho de que los tonos que componen cualquier sonido son armónicos del tono fundamental, es decir, que sus frecuencias son múltiplos enteros de la frecuencia fundamental. Pero Euler, en su *Tentamen Novae Theoriae Musicae* (1739) [16], mantenía que únicamente en instrumentos musicales apropiados los tonos superiores eran armónicos del tono fundamental. Se propuso por ello demostrar que la cuerda de espesor variable o de densidad $\sigma(x)$ y tensión T no uniformes emite tonos superiores no armónicos.

La ecuación diferencial resultantes

$$\frac{1}{c^2}\,\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}, \tag{28}$$

donde c es ahora función de x. Los primeros resultados interesantes los obtuvo Euler en el artículo «Sobre el movimiento vibratorio de cuerdas de espesor no uniforme» [17]. Afirma que la solución general no está al alcance de la potencia del análisis y obtiene una solución

[16] *Opera*, (3), 1, 197-427.

[17] *Novi Comm. Acad. Sci. Petrop.*, 9, 1762/63, 246-304, pub. 1764 = *Opera*, (2), 10, 293-343.

en el caso especial en que la distribución de masa σ está dada por

$$\sigma = \frac{\sigma_0}{\left(1 + \dfrac{x}{\alpha}\right)^4},$$

donde σ_0 y σ son constantes. Se tiene entonces

$$y = \left(1 + \frac{x}{\alpha}\right)\left[\phi\left(\frac{x}{1 + \dfrac{x}{\alpha}} + c_0 t\right) + \psi\left(\frac{x}{1 + \dfrac{x}{\alpha}} - c_0 t\right)\right],$$

donde $c_0 = \sqrt{T/\sigma_0}$. Las frecuencias de los modos o armónicos están dadas por

$$\nu_k = \frac{k}{2l}\left(1 + \frac{l}{\alpha}\right)\sqrt{T/\sigma_0}, \qquad k = 1, 2, 3, \ldots .$$

Por tanto, la razón de dos frecuencias sucesivas es la misma que para una cuerda de espesor uniforme, pero la frecuencia fundamental ya no es inversamente proporcional a la longitud.

En este artículo de 1762-63. Euler consideró también las vibraciones de una cuerda compuesta de dos trozos de longitudes a y b y grosores diferentes m y n. Dedujo la ecuación para las frecuencias ω de los modos, a saber,

$$m \operatorname{tg} \frac{\omega a}{m} + n \operatorname{tg} \frac{\omega b}{n} = 0, \tag{29}$$

resolviéndola en casos especiales. Las soluciones de (29) se denominan valores característicos o autovalores del problema. Estos valores son, como veremos, de capital importancia en la teoría de ecuaciones en derivadas parciales. Es casi evidente de (29) que las frecuencias características no son múltiplos enteros de la fundamental.

Pero Euler abordó de nuevo esta cuestión en otro artículo sobre la cuerda vibrante de espesor variable [18], y partiendo de (28) demuestra que existen funciones $c(x)$ para las que las frecuencias de los modos superiores no son múltiplos enteros de la fundamental.

---

[18] *Misc. Taur.*, 3, 1762/65, 25-29, pub. 1766 = *Opera*, (2), 10, 397-425.

También D'Alembert estudió la cuerda de espesor variable [19], utilizando un importante método de resolución que había introducido anteriormente para la cuerda de densidad constante. Se basa en la idea de separación de variables y hoy constituye un método fundamental para resolver ecuaciones en derivadas parciales [20]. Para resolver

$$\frac{\partial^2 y(t, x)}{\partial t^2} = a^2 \frac{\partial^2 y(t, x)}{\partial x^2}$$

D'Alembert pone

$$y = h(t)\, g(x),$$

sustituye esto en la ecuación diferencial y obtiene

$$\frac{1}{a^2} \frac{h''(t)}{h(t)} = \frac{g''(x)}{g(x)}. \qquad (30)$$

Razona después, como lo hacemos hoy, que como g''/g no varía con t, ha de ser una constante, y por el mismo argumento aplicado a h''/h, esta expresión también ha de ser una constante. Las dos constantes serán iguales y se denotan por A. Llega así a dos ecuaciones diferenciales ordinarias separadas

$$\begin{aligned} h''(t) - a^2 A h(t) &= 0 \\ g''(x) - A g(x) &= 0. \end{aligned} \qquad (31)$$

Como a y A son constantes, estas dos ecuaciones se resuelven fácilmente y D'Alembert obtiene

$$y(t, x) = h(t)\, g(x) = [Me^{a\sqrt{A}\cdot t} + Ne^{-a\sqrt{A}\cdot t}]\, [Pe^{a\sqrt{A}\cdot x} + Qe^{-\sqrt{A}\cdot t}].$$

Las condiciones de contorno, $y(t, 0) = y(t, l) = 0$, le llevaron a la conclusión de que $g(x)$ tenía que ser de la forma k sen Rx y que lo mismo ocurría con $h(t)$ ya que $y(t, x)$ había de ser periódica en t. D'Alembert dejó la cuestión en este punto. Daniel Bernoulli había

---

[19] *Hist. de l'Acad. de Berlin*, 19, 1763, 242 y sigs., pub. 1770.

[20] *Hist. de l'Acad. de Berlin*, 6, 1750, 335-360, pub. 1752.

utilizado la idea de separación de variables en 1732 en su tratamiento de las vibraciones de una cadena suspendida por un extremo, pero D'Alembert fue más explícito, a pesar de que no completó la resolución del problema.

En su artículo de 1763, D'Alembert escribió la ecuación de ondas en la forma

$$\frac{\partial^2 y}{\partial t^2} = X(x)\,\frac{\partial^2 y}{\partial x^2}$$

y buscó soluciones del tipo

$$u = \xi(x)\cos\lambda\pi t,$$

obteniendo para ξ la ecuación

$$\frac{d^2\xi}{dx^2} = \frac{-\lambda^2\pi^2\,\xi}{X(x)}. \tag{32}$$

Luego, había que determinar ξ de modo que valiese 0 en los dos extremos de la cuerda. Mediante un análisis detallado, D'Alembert demostró que existen valores de λ para los que ξ satisface esa condición, pero no vio que hay infinitos valores para λ. La importancia de la investigación de D'Alembert estriba en que supuso otro paso en el camino hacia los problemas de contorno, o problemas de autovalores, en ecuaciones diferenciales ordinarias.

Otro de los problemas abordados por Euler fue el de las oscilaciones transversales de una cuerda horizontal continua *pesada*. En el artículo «Sobre el efecto modificador de su propio peso sobre el movimiento de las cuerdas» [21], Euler obtiene la ecuación diferencial

$$\frac{1}{c^2}\,y_{tt} = \frac{g}{c^2} + y_{xx}.$$

Para c constante y en el caso de extremos fijos en $x = 0$ y $x = l$, Euler obtiene que

$$y = -\frac{(1/2)\,gx\,(x-l)}{c^2} + \phi(ct+x) + \psi(ct-x).$$

[21] *Acta Acad. Sci. Petrop.*, 1, 1781, 178-190, pub. 1784 = *Opera*, (2), 11, 324-334, aunque data de 1774.

Así pues, los resultados son los mismos que para la cuerda «sin peso» (cuando se desprecia la fuerza de la gravedad), excepto que la oscilación tiene lugar alrededor de la figura parabólica de equilibrio

$$y = -\frac{(1/2)\, gx(x-l)}{c^2}.$$

Como veremos en seguida, Euler había introducido todas las funciones de Bessel de primera especie en un artículo sobre la membrana vibrante (ver también cap. 21, secs. 4 y 6) y en este artículo de 1781 indica que es posible expresar cualquier movimiento mediante una serie de funciones de Bessel (a pesar de haberse opuesto a la afirmación de Daniel Bernoulli, en el problema de la cuerda vibrante, de que cualquier función se podía representar mediante una serie de funciones trigonométricas).

Se publicaron hasta finales de siglo muchos otros artículos, de los que los anteriores son sólo una muestra, sobre la cuerda vibrante y la cadena colgante. Los autores continuaron sin estar de acuerdo, corrigiéndose unos a otros y cometiendo toda clase de errores al hacerlo, incluyendo contradicciones con lo que ellos mismos habían dicho e incluso probado. Formularon afirmaciones, argumentos y refutaciones sobre la base de razonamientos poco rigurosos y, a menudo de, simplemente, predilecciones y convicciones personales. Sus referencias a artículos para probar sus argumentos no demostraban lo que afirmaban y también recurrían al sarcasmo, la ironía, la invectiva y el autobombo. Mezclados con estos ataques estaban los acuerdos aparentes para buscar favores, particularmente con D'Alembert, que tenía una considerable influencia con Federico II de Prusia y como director de la Academia de Ciencias de Berlín.

Los problemas de ecuaciones en derivadas parciales de segundo orden descritos hasta aquí hacían intervenir una sola variable espacial además del tiempo. En el siglo XVIII no se fue mucho más allá de esto. En un artículo de 1759 [22], Euler abordó las vibraciones de una membrana rectangular, considerando así un medio bidimensional, y obtuvo para los desplazamientos verticales z de la misma la ecuación

$$\frac{1}{c^2}\,\frac{\partial^2 z}{\partial t^2} = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}, \qquad (33)$$

[22] *Novi Comm. Acad. Sci. Petrop.*, 10, 1764, 243-260 pub. 1766 = *Opera*, (2), 10, 344-359.

en donde x e y representan las coordenadas de un punto genérico de la membrana y c está determinado por la masa y la tensión. Euler probó soluciones del tipo.

$$z = v(x, y) \text{ sen } (\omega t + \alpha),$$

obteniendo

$$0 = \frac{\omega^2 v}{c^2} + \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}.$$

Esta ecuación tiene soluciones sinusoidales de la forma

$$v = \text{sen} \left(\frac{\beta x}{a} + B \right) \text{ sen } \left(\frac{\gamma x}{b} + C \right),$$

donde

$$\frac{\omega^2}{c^2} = \frac{\beta^2}{a^2} + \frac{\gamma^2}{b^2}.$$

Las dimensiones de la membrana son a y b, de modo que $0 \leq x \leq a$ y $0 \leq y \leq b$. Cuando la velocidad inicial es 0, B y C se pueden tomar 0. Si los lados están fijos, entonces $\beta = m\pi$ y $\gamma = n\pi$, donde m y n son enteros. Por tanto, como $\omega = 2\pi\nu$, siendo ν la frecuencia por segundo, obtiene inmediatamente que las frecuencias son

$$\nu = \frac{1}{2} c \sqrt{\frac{m^2 + n^2}{a^2 + b^2}}.$$

Considera Euler después una membrana circular y transforma (33) a coordenadas polares (una idea muy original), obteniendo

$$\frac{1}{c^2} \frac{\partial^2 z}{\partial t^2} = \frac{\partial^2 z}{\partial r^2} + \frac{1}{r} \frac{\partial z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 z}{\partial \phi^2}. \tag{34}$$

Ensaya ahora soluciones de la forma

$$z = u(r) \text{ sen } (\omega t + A) \text{ sen } (\beta\phi + B), \tag{35}$$

obteniendo que $u(r)$ ha de satisfacer

$$u'' + \frac{1}{r} u'' + \left(\frac{\omega^2}{c^2} - \frac{\beta^2}{r^2} \right) u = 0. \tag{36}$$

Aparece aquí la ecuación de Bessel en la forma usual (cfr. cap. 21, sec. 6). Euler calcula entonces una solución en serie de potencias

$$u\left(\frac{\omega}{c}\ r\right) = r^{\beta}\left\{1 - \frac{1}{1(\beta+1)}\left(\frac{\omega}{c}\ \frac{r}{2}\right)^{2} + \frac{1}{1\cdot 2(\beta+1)\ (\beta+2)}\left(\frac{\omega}{c}\ \frac{r}{2}\right)^{4} + \cdots\right\},$$

que hoy escribiríamos como

$$u = \left(\frac{\omega}{c}\ r\right) = \left(\frac{c}{\omega}\right)^{\beta} 2^{\beta}\ \Gamma(\beta+\ 1)\, J_{\beta}\left(\frac{\omega}{c} r\right).$$

Como el borde $r=a$ ha de permanecer fijo,

$$J_{\beta}\left(\frac{\omega}{c}\ a\right) = 0. \tag{37}$$

También se sigue de (35), por ser z de período 2π respecto a ϕ, que β es un entero. Euler afirma que, para un β fijado, existen infinitas raíces de ω, de modo que resultarán infinitos sonidos simples; no calculó, sin embargo, esas raíces. Sí intentó hallar una segunda solución de (36), pero no lo logró. La teoría de la membrana vibrante fue desarrollada independientemente por Poisson [23] y a menudo se le atribuye a él únicamente.

Euler, Lagrange y otros trabajaron sobre la propagación del sonido en el aire. Euler escribió frecuentemente sobre el tema del sonido desde sus veinte años (1727), estableciendo este campo como una rama de la física matemática. Su mejor trabajo en la materia siguió a sus principales artículos sobre hidrodinámica de los años cincuenta del siglo. El aire es un fluido compresible y la teoría de la propagación del sonido es parte de la mecánica de fluidos (y de la elasticidad, pues el aire es también un medio elástico). No obstante, Euler hizo, para tratar la propagación del sonido, simplificaciones razonables de las ecuaciones generales de la hidrodinámica.

En 1759, se leyeron en la Academia de Berlín tres artículos excelentes y definitivos. En el primero, «Sobre la propagación del so-

---

[23] *Mém. de l'Acad. des Sci., Paris*, (2), 8, 1829, 357-570.

nido»[24], Euler considera la propagación del sonido en una dimensión espacial. Después de algunas aproximaciones, que vienen a significar el considerar ondas de pequeña amplitud, llega a la ecuación de ondas unidimensional

$$\frac{\partial^2 y}{\partial t^2} = 2gh \frac{\partial^2 y}{\partial x^2},$$

donde y es la amplitud de la onda en el punto x y el instante t, g es la aceleración de la gravedad y h es una constante que relaciona la presión y la densidad. Esta ecuación, como por supuesto vio Euler, es la misma que la de la cuerda vibrante y no hizo matemáticamente nada nuevo al resolverla.

En su segundo artículo[25], Euler da la ecuación bidimensional de propagación en la forma

$$\begin{aligned} \frac{\partial^2 x}{\partial t^2} &= c^2 \frac{\partial^2 x}{\partial X^2} + c^2 \frac{\partial^2 y}{\partial X\, \partial Y}, \\ \frac{\partial^2 y}{\partial t^2} &= c^2 \frac{\partial^2 x}{\partial Y^2} + c^2 \frac{\partial^2 x}{\partial X\, \partial Y}, \end{aligned} \tag{38}$$

donde x e y son las amplitudes de onda en las direcciones X e Y, respectivamente, o componentes del desplazamiento, y $c = \sqrt{2gh}$. Da la solución de tipo onda plana

$$x = \alpha\phi(\alpha X + \beta Y + c\sqrt{a^2 + \beta^2}\, t),\ y = \beta\phi(\alpha X + \beta Y + c\sqrt{\alpha^2 + \beta^2}\, t),$$

donde ϕ es una función arbitraria y α y β son constantes arbitrarias. Poniendo entonces

$$v = \frac{\partial x}{\partial X} + \frac{\partial y}{\partial Y} \tag{39}$$

(v se denomina divergencia del desplazamiento), Euler llega a la ecuación de ondas bidimensional

$$\frac{1}{c^2} \frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial X^2} + \frac{\partial^2 v}{\partial Y^2}. \tag{40}$$

[24] *Mém. de l'Acad de Berlin*, 15, 1759, 185-209, pub. 1776 = *Opera*, (3), 1, 428-451.
[25] *Mém. de l'Acad de Berlin*, 15, 1759, 210-240, pub. 1776 = *Opera*, (3), 1, 452-483.

Afirma aquí también la necesidad de superponer soluciones para obtener la solución más general posible, a fin de satisfacer una condición inicial dada, es decir, el valor de v o de x e y en $t = 0$.

Euler muestra a continuación cómo conseguir la ecuación diferencial cuyas soluciones se denominan ondas cilíndricas, por propagarse como cilindros que se expanden. Pone $Z = \sqrt{X^2 + Y^2}$ e introduce $v = f(X, t)$, donde f es arbitraria. Poniendo $x = vX$ e $y = vY$, obtiene de (40)

$$\frac{1}{c^2}\,\frac{\partial^2 v}{\partial t^2} = \frac{3}{Z}\,\frac{\partial v}{\partial Z} + \frac{\partial^2 v}{\partial Z^2}.$$

También obtiene en este artículo, de una manera análoga, la ecuación de ondas tridimensional

$$\frac{1}{c^2}\,\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial X^2} + \frac{\partial^2 v}{\partial Y^2} + \frac{\partial^2 v}{\partial Z^2}, \qquad (41)$$

donde v es de nuevo la divergencia del desplazamiento (x, y, z). Euler da soluciones de tipo ondas planas y ondas esféricas utilizando sustituciones análogas a la indicada para las ondas cilíndricas. La ecuación fundamental para las ondas esféricas es

$$\frac{1}{c^2}\,\frac{\partial^2 s}{\partial t^2} = \frac{4}{V}\,\frac{\partial s}{\partial V} + \frac{\partial^2 s}{\partial V^2}, \qquad V = \sqrt{X^2 + Y^2 + Z^2}.$$

Buena parte de este trabajo sobre ondas cilíndricas y esféricas fue también realizado independientemente por Lagrange a finales del año 1759. Se comunicaban los resultados uno al otro, y aunque hay muchos detalles en los que el trabajo de Lagrange difiere del de Euler, no hay cuestiones matemáticas importantes que merezcan reseñarse aquí.

De la propagación de ondas de sonido en el aire sólo había un paso al estudio de los sonidos emitidos por instrumentos musicales que utilizan el movimiento del aire. Este estudio fue iniciado por Daniel Bernoulli en 1739. Bernoulli, Euler y Lagrange escribieron numerosos artículos sobre los tonos emitidos por una variedad casi increíble de tales instrumentos. En una publicación de 1762, Daniel Bernoulli demostró que no puede darse condensación de aire en el

extremo abierto de un tubo cilíndrico (un tubo de órgano)[26]. En un extremo cerrado, las partículas de aire han de estar en reposo. Dedujo de ello que un tubo cerrado o abierto en ambos extremos tiene el mismo modo fundamental que un tubo la mitad de largo, pero abierto en un extremo y cerrado en el otro. Descubrió también el teorema de que para tubos de órgano cerrados las frecuencias de los armónicos son múltiplos impares de la frecuencia de tono fundamental. En el mismo artículo, Bernoulli consideró tubos de forma distinta a la cilíndrica, en particular el tubo cónico, para el que obtuvo expresiones para los tonos (modos) individuales pero reconoció que sólo eran válidas para conos infinitos y no para el cono truncado. Para el tubo cónico (infinito), los tonos superiores resultaron ser armónicos del fundamental. Bernoulli confirmó muchos de sus resultados teóricos mediante experimentos.

También estudió Euler tubos cilíndricos y figuras no cilíndricas de revolución[27] y consideró la reflexión en extremos cerrados y abiertos. Los esfuerzos de estos pensadores estaban dirigidos a comprender las flautas, tubos de órgano, toda clase de trompas con forma de cilindro, de cono y de hiperboloide, trompetas, cornetas y otros instrumentos de viento.

En su conjunto, estos esfuerzos para resolver ecuaciones en derivadas parciales en tres y cuatro variables fueron limitados, principalmente porque las soluciones se expresaban mediante series en que aparecían varias variables, en contraste con la mayor simplicidad de las series trigonométricas en x y t separadamente (cfr. [20]). Pero los matemáticos sabían poco de las funciones que aparecían en esas series más complicadas y de los métodos para determinar los coeficientes, métodos que pronto se desarrollarían.

Merece la pena mencionar que Euler, al considerar el sonido de una campana y al reconsiderar algunos de los problemas de la vibración de barras, hubo de plantearse ecuaciones en derivadas parciales de cuarto orden. Sin embargo, no fue capaz de hacer mucho con ellas y, de hecho, su estudio no avanzó en lo que quedaba de siglo.

---

[26] *Mém. de l'Acad. des Sci., Paris*, 1762, 431-485, pub. 1764.

[27] *Novi Comm. Acad. Sci. Petrop.*, 16, 1771, 281-425, pub. 1772 = *Opera*, (2), 13, 262-369.

4. Teoría del potencial

El desarrollo de las ecuaciones en derivadas parciales fue impulsado además por otra clase de investigaciones físicas. Uno de los principales problemas del siglo XVIII era la determinación de la magnitud de la atracción gravitatoria que una masa ejerce sobre otra, siendo los principales casos el de la atracción del Sol sobre un planeta, la de la Tierra sobre una partícula exterior o interior a ella y la de la Tierra sobre otra masa extensa. Cuando dos masas están muy alejadas en comparación con sus tamaños, cabe tratarlas como masas puntuales; pero en otros casos, señaladamente el de la Tierra atrayendo una partícula, hay que tener en cuenta el tamaño de la Tierra. Claramente, hay que conocer la forma de la Tierra para calcular la atracción gravitatoria que su masa distribuida ejerce sobre una partícula o sobre otra masa distribuida. Aunque la forma precisa continuó siendo un tema de investigación (cap. 21, sec. 1), era claro en 1700 que debería ser algún tipo de elipsoide, quizá un esferoide achatado (un elipsoide engendrado al girar una elipse alrededor del eje menor). Para un esferoide achatado sólido, la fuerza de atracción, lo mismo sobre una partícula externa que sobre una interna, no se puede calcular como si la masa estuviese concentrada en el centro.

En un artículo premiado de 1740 sobre las mareas, y en su *Treatise of Fluxions* (1742), Maclaurin demostró que, para un fluido de densidad uniforme y bajo rotación angular constante, el esferoide achatado es una figura de equilibrio. A continuación, Maclaurin demostró sintéticamente que dados dos elipsoides homogéneos de revolución homofocales, las atracciones que ambos cuerpos ejercen sobre una misma partícula exterior a ambos, supuesto que la partícula esté en la prolongación del eje de revolución o en el plano del ecuador, serán proporcionales a los volúmenes. En el siglo XIX, James Ivory (1765-1842) y Michel Chasles también establecieron geométricamente algunos otros resultados particulares.

El enfoque geométrico del problema de la atracción gravitatoria utilizado por Newton, Maclaurin y otros es apropiado sólo para cuerpos especiales y para posiciones especiales de las masas bajo atracción. Dicho enfoque pronto dio paso a los métodos analíticos, los cuales aparecen en primer lugar en artículos de Clairaut antes de 1743 y especialmente en su célebre libro *Théorie de la figure de la Terre* (1743), en el cual considera tanto la forma de la Tierra como la atracción gravitatoria.

Señalemos en primer lugar algunos hechos relativos a la formulación analítica. La fuerza de gravitación ejercida por un cuerpo extenso sobre una unidad de masa P considerada como una partícula es la suma de las fuerzas ejercidas por todas las pequeñas masas que constituyen el cuerpo. Si $d\xi\, d\eta\, d\zeta$ representa un pequeño volumen del cuerpo (fig. 22.2), tan pequeño como para poder considerarlo localizado en el punto (ξ, η, ζ), y si P tiene por coordenadas (x, y, z), la atracción ejercida por la pequeña masa de densidad ϱ sobre la partícula unidad es un vector dirigido de P a la pequeña masa que, por la ley de gravitación de Newton, tiene por componentes (cap. 21, sec. 7)

$$-k\varrho\,\frac{x-\xi}{r^3}\,d\xi\,d\eta\,d\zeta, \qquad -k\varrho\,\frac{y-\eta}{r^3}\,d\xi\,d\eta\,d\zeta,$$

$$-k\varrho\,\frac{z-\zeta}{r^3}\,d\xi\,d\eta\,d\zeta,$$

donde k es la constante en la ley de Newton y

$$r=\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}.$$

Naturalmente, ϱ puede ser función de ξ, η y ζ o, en el caso de un cuerpo homogéneo, constante.

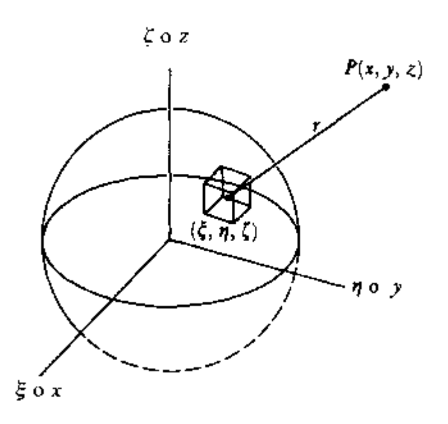


FIGURA 22.2

La fuerza que ejerce todo el cuerpo sobre la masa unidad situada en P tiene por componentes:

$$f_x = -k \iiint \varrho \frac{x-\xi}{r^3} \, d\xi \, d\eta \, d\zeta$$

$$f_y = -k \iiint \varrho \frac{y-\eta}{r^3} \, d\xi \, d\eta \, d\zeta \qquad (42)$$

$$f_z = -k \iiint \varrho \frac{z-\zeta}{r^3} \, d\xi \, d\eta \, d\zeta$$

en donde la integral se extiende sobre todo el cuerpo atractor. Estas integrales son finitas y válidas también cuando el punto P está en el interior del cuerpo atractor.

En lugra de tratar separadamente cada componente de la fuerza, es posible introducir una función $V(x, y, z)$ cuyas derivadas parciales respecto a x, y y z sean respectivamente las tres componentes de la fuerza. Esta función es

$$V(x, y, z) = \iiint \frac{\varrho}{r} \, d\xi \, d\eta \, d\zeta. \qquad (43)$$

Derivando bajo el signo integral respecto a x, y, z (incluidas en r) se obtiene

$$\frac{\partial V}{\partial x} = \frac{1}{k} f_x, \qquad \frac{\partial V}{\partial y} = \frac{1}{k} f_y, \qquad \frac{\partial V}{\partial z} = \frac{1}{k} f_z,$$

siendo también válidas estas ecuaciones cuando P está en el interior del cuerpo atractor. La función V se dice que es una función potencial. Cuando problemas que involucran las tres componentes f_x, f_y, f_z se pueden reducir a trabajar con V, se tiene la ventaja de trabajar con una sola función en lugar de con tres.

Si se conoce la distribución de la masa en el interior del cuerpo, lo que significa conocer ϱ como función de ξ, η y ζ, y si se conoce la forma exacta del cuerpo, se puede en ocasiones obtener V mediante el cálculo efectivo de la integral. Sin embargo, esa integral triple no es expresable en términos de funciones simples para la

mayor parte de formas de cuerpos y además no se conoce la distribución real de la masa en el interior de la Tierra y de otros cuerpos. En consecuencia, V se ha de calcular por otros métodos. El hecho principal relativo a V es que para puntos (x, y, z) *exteriores* al cuerpo atractor dicha función satisface la ecuación en derivadas parciales

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{44}$$

en la que se advierte que ya no aparece ϱ. Esta ecuación diferencial se conoce como ecuación del potencial o ecuación de Laplace.

La idea de que una fuerza puede derivar de una función potencial, e incluso el término de «función potencial», fueron utilizados por Daniel Bernoulli en su *Hydrodynamica* (1738). La propia ecuación del potencial aparece por primera vez en uno de los principales artículos de Euler, elaborado en 1752, «Principios del movimiento de fluidos» [28]. Al estudiar las componentes u, v y w de la velocidad de un punto del fluido, Euler había probado que $u\,dx + v\,dy + w\,dz$ tenía que ser una diferencial exacta. Introduce la función S tal que $dS = u\,dx + v\,dy + w\,dz$ y entonces

$$u = \frac{\partial S}{\partial x}, \qquad v = \frac{\partial S}{\partial y}, \qquad w = \frac{\partial S}{\partial z}.$$

Pero el movimiento de los fluidos incompresibles obedece a la llamada ley de continuidad, a saber,

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{45}$$

que expresa matemáticamente el hecho de que no se crea ni se destruye materia durante el movimiento. Se tiene entonces

$$\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} + \frac{\partial^2 S}{\partial z^2} = 0.$$

Euler dice que no se conoce cómo resolver esta ecuación en general, por lo que considera sólo casos especiales en los que S es un poli-

[28] *Novi Comm. Acad. Sci. Petrop.*, 6, 1756/67, 271-311, pub. 1761 = *Opera*, (2), 12, 133-168.

nomio en x, y, z. La función S fue denominada más tarde (1868) potencial de velocidades por Helmholtz. En un artículo publicado en 1762 [29], Lagrange reprodujo todas estas cantidades, que él tomó de Euler sin mencionarlo aunque sí mejoró la exposición en cuanto a ideas y expresiones.

Antes de que podamos examinar los trabajos realizados para resolver la ecuación del potencial y aplicarla a la atracción gravitatoria, debemos pasar revista a algunos esfuerzos para evaluar esa atracción directamente por medio de las integrales (42) o sus equivalentes en otros sistemas coordenados.

En un artículo escrito en 1782 aunque publicado en 1785 y titulado «Recherches sur l'attraction des sphéroïdes» [30], Legendre, interesado en la atracción ejercida por sólidos de revolución, demostró el siguiente teorema: si se conoce la atracción de un sólido de revolución sobre todo punto exterior situado en la prolongación de su eje, entonces se conoce para todo punto exterior. Primero, expresó la componente de la fuerza de atracción en la dirección del radio vector r por medio de

$$P(r, \theta, 0) = \iiint \frac{(r - r') \cos \gamma}{(r^2 - 2rr' \cos \gamma + r'^2)^{3/2}} r'^2 \operatorname{sen} \theta' \, d\theta' \, d\phi' \, dr', \quad (46)$$

donde (fig. 22.3) r es el radio vector hasta el punto sujeto a la atracción, r' es el radio vector hasta un punto genérico del cuerpo que

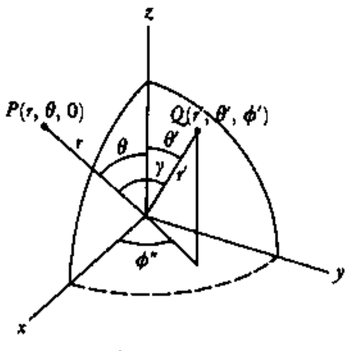


FIGURA 22.3

[29] *Misc. Taur.*, 2_2 1760/61, 196-298, pub. 1762 = *Œuvres*, 1, 365-468.
[30] *Mém. des sav. étrangers*, 10, 1785, 411-434.

ejerce la atracción y γ es el ángulo formado en el centro del cuerpo por los dos radios vectores. La coordenada ϕ del punto exterior puede tomarse 0 porque el sólido es una figura de revolución alrededor del eje z. Legendre desarrolla después el integrando en potencias de r'/r. Esto se hace escribiendo el denominador como

$$r^3\left[1-\left(2\,\frac{r'}{r^2}\cos\gamma-\frac{r'^2}{r^2}\right)\right]^{3/2}.$$

La cantidad entre corchetes se puede poner en el numerador y desarrollarse después por el teorema del binomio con la cantidad entre paréntesis como segundo término del binomio. Legendre obtuvo para el integrando, aparte el elemento de volumen, la serie

$$\frac{1}{r^2}\left\{1+3P_2(\cos\gamma)\,\frac{r'^2}{r^2}+5P_4(\cos\gamma)\,\frac{r'^4}{r^4}\right.$$
$$\left.+\,7P_6(\cos\gamma)\,\frac{r'^6}{r^6}+\cdots\right\}.$$

Los coeficientes $P_2, P_4, \ldots$ son funciones racionales enteras de cos γ; son las funciones que hoy se conocen como polinomios de Legendre (o coeficientes o armónicos zonales de Laplace). Legendre dio la forma que tienen de modo que se pudiese deducir el P_n general, a saber,

$$P_n(x)=\frac{(2n-1)\,(2n-3)\quad\cdots\quad 1}{n!}\cdot$$
$$\left[x^n-\frac{n(n-1)}{2(2n-1)}\,x^{n-2}\right.$$
$$\left.+\,\frac{n(n-1)\,(n-2)\,(n-3)}{2\cdot 4\cdot(2n-1)\,(2n-3)}\,x^{n-4}+\cdots\right]. \qquad (47)$$

Podía así integrar en r' para obtener

$$\frac{2}{r^2}\iint\left\{\frac{R^3}{3}+\frac{3}{5}\,P_2(\cos\gamma)\,\frac{R^5}{r^2}\right.$$
$$\left.+\,\frac{5}{7}\,P_4(\cos\gamma)\,\frac{R^7}{r^4}+\cdots\right\}\ \text{sen}\ \theta'\,d\theta'\,d\phi',$$

donde $R = f(\theta')$ es el valor de r' en un θ' dado (es independiente de ϕ'). A continuación tenía que integrar respecto a θ'; para ello utilizó [31]

$$\cos \gamma = \cos \theta \cos \theta' + \operatorname{sen} \theta \operatorname{sen} \theta' \cos \phi'.$$

Después de establecer el resultado auxiliar

$$\frac{1}{\pi} \int_0^{\pi} P_{2n} (\cos \gamma)\, d\phi' = P_{2n} (\cos \theta)\, P_{2n} (\cos \theta'),$$

obtuvo finalmente

$$P(r, \theta, 0) = \frac{3M}{r^2} \sum_{n=0}^{\infty} \frac{2n-3}{2n-1} P_{2n} (\cos \theta) \frac{\alpha_n}{r^{2n}},$$

donde

$$\alpha_n = \frac{4\pi}{3M} \int_0^{\pi/2} R^{2n+3}\, P_{2n}(\cos \theta') \operatorname{sen} \theta'\, d\theta'.$$

El valor de esta integral depende de la forma de las curvas meridianas $R = f(\theta')$.

Del resultado anterior, y gracias a una comunicación recibida de Laplace, Legendre obtuvo la expresión de la función potencial, y de ésta dedujo la componente de la fuerza de atracción perpendicular al radio vector.

En un segundo artículo escrito en 1784 [32], Legendre dedujo algunas propiedades de las funciones P_{2n}. Así,

$$\int_0^1 f(x^2)\, P_{2n}(x)\, dx = 0 \tag{48}$$

---

[31] Esta expresión se deduce como sigue: de acuerdo con las ecuaciones que dan la transformación de coordenadas esféricas a coordenadas rectangulares, $x = r \operatorname{sen} \theta$ con ϕ, $y = r \operatorname{sen} \theta \operatorname{sen} \phi$, $z = r \cos \theta$, las coordenadas rectangulares de P son ($r \operatorname{sen} \theta \cos \phi$, $r \operatorname{sen} \theta \operatorname{sen} \phi$, $r \cos \theta$) y las de Q ($r' \operatorname{sen} \theta' \cos \phi'$, $r' \operatorname{sen} \theta' \operatorname{sen} \phi$, c' xoa θ'). Utilizando entonces la fórmula de la distancia, podremos expresar PQ. Pero, por la ley de los cosenos, $PQ = r^2 + r'^2 + 2rr' \cos \gamma$. Igualando las dos expresiones para PQ se tendrá la expresión para $\cos \gamma$ que se utiliza en el texto.

[32] *Mém. de l'Acad. des Sci., Paris*, 1784, 370-389, pub. 1787.

para toda función racional entera de x^2 cuyo grado en x^2 sea menor que n. Si n es un entero positivo cualquiera,

$$\int_0^1 x^n P_{2m}\, dx = \frac{n(n-2)\cdots(n-2m+2)}{(n+1)(n+3)\cdots(n+2m+1)}. \qquad (49)$$

Si m y n son enteros positivos

$$\int_0^1 P_{2n}(x) P_{2m}(x)\, dx = \begin{cases} 0 \text{ para } m \gtrless n, \\ \dfrac{1}{4m+1} \text{ para } m = n. \end{cases} \qquad (50)$$

Demostró también que los ceros de cada uno de los P_{2n} son reales, distintos entre sí, simétricos respecto a 0 y de valor absoluto menor que 1. Asimismo, $P_{2n}(x) < 1$ para $0 < x < 1$.

Después, con ayuda de la condición de ortogonalidad (50), demuestra (por integración de la serie término a término) que una función dada de x^2 se puede expresar de modo único mediante una serie de funciones $P_{2n}(x)$.

Finalmente, utilizando éstas y otras propiedades de sus polinomios, Legendre vuelve al problema principal de la atracción gravitatoria y utilizando la expresión (43) para el potencial y la condición de equilibrio de una masa de fluido que gira, obtiene la ecuación de la curva meridiana de tal masa en la forma de una serie de sus polinomios. Legendre creyó que esa ecuación incluía todas las posibles figuras de equilibrio para un esferoide de revolución.

Entra entonces en esta historia Laplace, quien había escrito varios artículos sobre la fuerza de atracción ejercida por volúmenes de revolución (1772, pub. 1776; 1773, pub. 1776, y 1775, pub. 1778), en los cuales había trabajado con las componentes de la fuerza pero no con la función potencial. El artículo de Legendre de 1782, publicado en 1785, inspiró un célebre y destacado cuarto artículo de Laplace, «Théorie des attractions des sphéroïdes et de la figure des planètes» [33]. Sin mencionar a Legendre, Laplace abordó el problema de la atracción ejercida por un esferoide arbitrario en contraste con las figuras de revolución de Legendre. Por esferoide, Laplace entendía cualquier superficie dada por una ecuación en r, θ y ϕ.

Comienza con el teorema que establece que el potencial V de la

[33] *Mém. de l'Acad. des Sci., Paris*, 1782, 113-196, pub. 1785 = Œuvres, 10, 339-419.

fuerza que un cuerpo arbitrario ejerce sobre un punto exterior, expresado en coordenadas esféricas r, θ, ϕ con $\mu = \cos\theta$, satisface la ecuación del potencial

$$\frac{\partial}{\partial\mu}\left((1-\mu^2)\frac{\partial V}{\partial\mu}\right) + \frac{1}{1-\mu^2}\frac{\partial^2 V}{\partial\phi^2} + r\frac{\partial^2(rV)}{\partial r^2} = 0. \tag{51}$$

Laplace no dice aquí cómo obtuvo la ecuación. En un artículo posterior [34] da la ecuación en coordenadas rectangulares (44); lo más seguro es que dispusiese primero de la forma rectangular y que derivase de ella la forma en coordenadas esféricas. De hecho, ambas formas habían sido ya dadas por Euler y Lagrange, pero Laplace no los menciona; puede ser que no conociese su trabajo, aunque es dudoso.

En el artículo de 1782 Laplace pone

$$V(r, \theta, \phi) = \frac{U_0}{r} + \frac{U_1}{r^2} + \frac{U_2}{r^3} + \cdots, \tag{52}$$

donde $U_n = U_n(\theta, \phi)$, y sustituye esto en (51). Resulta entonces que cada U_n satisface [35]

$$\frac{\partial}{\partial\mu}\left((1-\mu^2)\frac{\partial U}{\partial\mu}\right) + \frac{1}{1-\mu^2}\frac{\partial^2 U}{\partial\phi^2} + n(n+1)\,U = 0. \tag{53}$$

Con ayuda de los polinomios P_{2n} de Legendre consigue demostrar que

$$U_n(\theta, \phi) = \iiint r'^{n+2}\,P_{2n}(\cos\theta\cos\theta' + \\ + \operatorname{sen}\theta\operatorname{sen}\theta'\cos(\phi-\phi'))\cdot\operatorname{sen}\theta'\,d\theta'\,d\phi'\,dr'. \tag{54}$$

---

[34] *Mém. de l'Acad. des Sci., Paris*, 1787, 249-267, pub. 1789 = Œuvres, 11, 275-292.

[35] Si despreciamos el término intermedio (o sea, no interviene ϕ), la ecuación que resulta es la que hoy conocemos como ecuación diferencial de Legendre,

$$(1-x^2)\frac{d^2z}{dx^2} - 2x\frac{dz}{dx} + n(n+1)\,z = 0.$$

Los $P_n(x)$ satisfacen esta ecuación. Por otro lado, las U_n (y las Y_n de [57]) consideradas como funciones de las dos variables $\mu = \cos\phi$ y ϕ satisfacen (53). Los alemanes llamaron a las U_n y Y_n funciones esféricas, mientras que lord Kelvin las llamó armónicos esféricos o armónicos esféricos de superficie.

A continuación, Laplace utiliza este resultado y (52) para calcular el potencial de un esferoide que difiere poco de una esfera. Escribe la ecuación de la superficie del esferoide como

$$r = a(1 + \alpha y), \tag{55}$$

donde α es pequeño e y es una función de θ' y ϕ' sobre el esferoide. Laplace supone que $y(\theta, \phi)$ se puede desarrollar en serie de funciones

$$y = Y_0 + Y_1 + Y_2 + \cdots, \tag{56}$$

donde las Y_n son funciones de θ y ϕ y satisfacen la ecuación diferencial (53). Obtiene en primer lugar el resultado de que

$$Y_n = \frac{2n+1}{4\pi\alpha a^{n+1}} U_n. \tag{57}$$

Estas Y_n, entonces, se pueden utilizar en (52). Asimismo, el desarrollo (56) se puede escribir en la forma

$$y(\mu, \phi) = \frac{1}{4\pi} \sum_{n=0}^{\infty} (2n+1)$$

$$\int_{-1}^{1} \int_{0}^{2\pi} Y_n(\mu', \phi') P_n(\mu, \phi, \mu', \phi') \, d\mu' \, d\phi', \tag{58}$$

donde $\mu' = \cos \theta'$. Con el valor de y se tiene por (55) el de r y con éste y los U_n dados por (54) se obtiene V de (52).

Laplace no considera aquí el problema general del desarrollo de *cualquier* función de θ y ϕ en serie de las Y_n. Supone, aquí como en artículos posteriores, que tal desarrollo es posible y que es único. Maneja funciones racionales enteras de μ, $\sqrt{1-\mu^2} \cos \phi$ y $\sqrt{1-\mu^2} \operatorname{sen} \phi$, con lo que no le es tan necesario el resultado general. Sí demuestra la propiedad fundamental de ortogonalidad

$$\int_{-1}^{1} \int_{0}^{2\pi} U_n(\mu, \phi) \, U_m(\mu, \phi) = 0, \qquad m \neq n. \tag{59}$$

No obstante, en su *Mécanique céleste*, volumen 2, demuestra que una función arbitraria de θ y ϕ se puede desarrollar en serie de las U_n(o las Y_n) y que (59) implica la unicidad de tal desarrollo.

Laplace escribió varios artículos más sobre la atracción de esferoides y sobre la forma de la Tierra (por ejemplo, 1783, pub. 1786; 1787, pub. 1789) y en esos artículos utiliza desarrollos en funciones esféricas. En el último artículo, en el que Laplace dio la ecuación del potencial en coordenadas rectangulares, cometió un error de importancia al suponer que dicha ecuación es válida cuando la masa puntual atraída por el cuerpo está en el interior de éste. Este error fue corregido por Poisson (cap. 28, sec. 4).

Legendre continuó sus investigaciones durante los años ochenta del siglo. Su cuarto artículo, escrito en 1790 [36], introdujo los $P_n(x)$ para *n impar*. (La expresión [47] para los $P_n(x)$ es válida para todo *n*.) Legendre demuestra que para todo par de enteros positivos *m* y *n* se tiene

$$\int_{-1}^{1} P_m(x)\, P_{2n}(x)\, dx = \begin{cases} 0 \text{ para } m \neq n, \\ \dfrac{2}{2n+1} \text{ para } m = n. \end{cases} \tag{60}$$

Después, también él introduce las funciones esféricas: hace que Y_n sea el coeficiente de z^n en el desarrollo de $(1-2zt+z^2)^{-1/2}$ con $t = \cos\theta\cos\theta' + \operatorname{sen}\theta\operatorname{sen}\theta'\cos(\phi-\phi')$ y, entonces, poniendo $\mu = \cos\theta$, $\mu' = \cos\theta'$ y $\psi = \phi - \phi'$, demuestra que

$$Y_n(t) = P_n(\mu)P_n(\mu') + \frac{2}{n(n+1)}\frac{dP_n(\mu)}{d\mu}\frac{dP_n(\mu')}{d\mu'}\operatorname{sen}\theta\operatorname{sen}\theta'\cos\psi$$

$$+\frac{2}{(n-1)(n+1)(n+2)}\frac{d^2P_n(\mu)}{d\mu^2}\frac{d^2P_n(\mu')}{d\mu'^2}\operatorname{sen}^2\theta\operatorname{sen}^2\theta'\cos 2\psi + \cdots,$$

conteniendo los términos siguientes derivadas superiores de P_n. Esta

[36] *Mém. de l'Acad. des Sci., Paris*, 1789, 372-454, pub. 1793.

ecuación es equivalente a

$$P_n(\cos\theta\cos\theta' + \operatorname{sen}\theta\operatorname{sen}\theta'\cos(\phi-\phi')) =$$

$$\sum_{m=1}^{n} P_n^m(\cos\theta)\,P_n^m(\cos\theta')\cos m(\phi-\phi'),$$

en donde m es un superíndice y no un exponente. Los $P_n^m(x)$ satisfacen

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_n^m}{dx}\right] + \left[n(n+1) - \frac{v^2}{1\times x^2}\right]P_n^m = 0 \qquad (61)$$

y, además, los $P_n^m(x)$ coinciden salvo un factor constante con

$$(1-x^2)^{m/2}\,\frac{d^m P_n^m(x)}{dx^m}.$$

Los $P_n^m(x)$ así introducidos se llaman ahora polinomios de Legendre asociados. Legendre demuestra después que

$$\int_{-1}^{1}\int_0^{2\pi} U_n(\mu',\phi')\,P_n(\mu,\phi,\mu',\phi')\,d\mu'\,d\phi' = \frac{4\pi}{2n+1}U_n(\mu,\phi) \qquad (62)$$

y que

$$\int_{-1}^{1}\int_0^{2\pi} \{P_n(\mu)\}^2\,d\mu\,d\phi = \frac{4\pi}{2n+1}. \qquad (63)$$

En este artículo se utiliza el hecho de que los $P_n(x)$ satisfacen la ecuación diferencial de Legendre.

Este, Laplace y otros obtuvieron muchos más resultados especiales relativos a los polinomios de Legendre y a los armónicos esféricos. Un resultados fundamental es la fórmula de Olinde Rodrigues (1794-1851), dada en 1816 [37],

$$P_n(x) = \frac{1}{2^n n!}\,\frac{d^n(x^2-1)^n}{dx^n}. \qquad (64)$$

[37] *Corresp. sur l'Ecole Poly.*, 3, 1816, 361-385.

El trabajo de Laplace sobre la solución de la ecuación del potencial para la fuerza de atracción de los esferoides significó el comienzo de un inmenso trabajo sobre el tema. Igualmente importante fue su trabajo, y el de Legendre, sobre los polinomios de Legendre $P_n(x)$, los polinomios de Legendre asociados $P_n^m(x)$ y los armónicos esféricos (de superficie) $Y_n(\mu, \phi)$, debido a que funciones con un alto grado de generalidad se pueden representar en serie de los P_n, los P_n^m y los Y_n. Estas series de funciones son análogas a las series trigonométricas de las que Bernoulli afirmaba que se podían utilizar para representar funciones arbitrarias. La elección de la clase de funciones depende de la ecuación diferencial a resolver y de las condiciones iniciales y de contorno. Por supuesto, quedaba mucho por hacer con estas funciones para que fuesen más útiles en la resolución de ecuaciones en derivadas parciales.

5. Ecuaciones en derivadas parciales de primer orden

Hasta la época de Lagrange se había llevado a cabo muy poco trabajo sistemático sobre las ecuaciones en derivadas parciales de primer orden. La atención se había centrado en las ecuaciones de segundo orden a causa de los problemas físicos que conducen directamente a ellas. Habían sido resueltas algunas ecuaciones de primer orden especiales, pero porque se integraban inmediatamente o mediante algún artificio ingenioso. Había una excepción, la que hoy se conoce habitualmente como ecuación en diferenciales totales y que tiene la forma

$$P\,dx + Q\,dx + R\,dz = 0, \tag{65}$$

donde P, Q y R son funciones de x, y, z. Si tal ecuación es integrable, define z como función de x e y. Clairaut se encontró con ecuaciones de este tipo en 1739 en su trabajo sobre la forma de la Tierra [38]. Si la expresión del primer miembro de (65) es una diferencial exacta, es decir, si existe una función $u(x, y, x)$ tal que

$$du = P\,dx + Q\,dy + R\,dz, \tag{66}$$

[38] *Mém. de l'Acad. des Sci., Paris*, 1739, 425-436.

entonces, señala Clairaut,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}. \tag{67}$$

Clairaut mostró cómo resolver (65) por un método que todavía se usa en los textos modernos. El interés de la ecuación (65) procede del hecho de que si P, Q y R son las componentes de la velocidad en el movimiento de un fluido, entonces (65) ha de ser una diferencial exacta.

Clairaut demostró también que si (65) no es exacta, entonces cabe que sea posible encontrar un factor integrante, es decir, una función $\mu(x, y, z)$ tal que al multiplicar por ella la ecuación el nuevo primer miembro resulte una diferencial exacta. Clairaut[39] y más tarde D'Alembert (*Traité de l'équilibre et du mouvement des fluides*, 1744) dieron una condición necesaria para que la ecuación sea integrable (con la ayuda de un factor integrante). Esta condición (que también es suficiente) es

$$P\left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right) + Q\left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) + R\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = 0. \tag{68}$$

La ecuación en derivadas parciales de primer orden en dos variables independientes *general* es de la forma

$$f(x, y, z, p, q) = 0, \tag{69}$$

donde $p = \partial z/\partial x$ y $q = \partial z/\partial y$. Si la ecuación es lineal, en p y q, entonces se trata de una ecuación en derivadas parciales lineal y si no será una ecuación no lineal. Lo más importante de la teoría fue aportado por Lagrange.

Para comprender el trabajo de Lagrange, hay que dar cuenta, antes que nada, de su terminología, la cual todavía está en uso. Lagrange clasificaba las soluciones de las ecuaciones de primer orden no lineales como sigue. Una solución $V(x, y, z, a, b) = 0$ que contenga dos constantes arbitrarias es la *solución completa* o *integral completa* de la ecuación. Haciendo $b = \phi(a)$, con ϕ función arbitraria, se obtiene una familia uniparamétrica de soluciones. Con $\phi(a)$

[39] *Mém. de l'Acad. des Sci., Paris*, 1740, 293-323.

arbitraria, la envolvente de esa familia es la *integral general* de la ecuación; si se considera una $\phi(a)$ concreta, la envolvente es un *caso particular* de la integral general. La envolvente de todas las soluciones incluidas en la integral completa se llama *integral singular* de la ecuación. Veremos después lo que significan geométricamente estas soluciones. La integral completa no es única en el sentido de que puede haber muchas diferentes que no se obtienen una de la otra mediante un simple cambio en las constante arbitrarias; pero de cualquiera de ellas se pueden obtener todas las soluciones dadas por otra por medio de los casos particulares y la integral singular.

Entre dos importantes artículos de 1772 y 1779, que examinaremos en seguida, Lagrange escribió un artículo en 1774 en el que discutía las relaciones entre las soluciones singular, general y completa de una ecuación en derivadas parciales de primer orden. La integral general se obtiene eliminando a entre $V(x, y, z, a, \phi(a)) = 0$ y $\partial V/\partial a = 0$, donde $\phi(a)$ es arbitraria [40]. (Para una $\phi(a)$ particular se obtiene una solución particular.) La solución singular se obtiene eliminando a y b entre $V(x, y, z, a, b) = 0$, $\partial V/\partial a = 0$ y $\partial V/\partial b = 0$.

Lagrange desarrolló primero la teoría general de las ecuaciones de primer orden *no lineales*. En el artículo de 1772 [41], considera una ecuación general de primer orden en dos variables independientes x e y, con z como variable dependiente, mejorando y generalizando lo que antes había hecho Euler. Considera la ecuación (69) dada en la forma

$$q - Q(x, y, z, p) = 0, \tag{70}$$

es decir, con q como función de x, y, z, p, y trata de determinar p como función de x, y, z de modo que las dos ecuaciones

$$q - Q(x, y, z, p) = 0 \quad \text{y} \quad p - P(x, y, z) = 0 \tag{71}$$

tengan una infinidad simple de superficies integrales comunes, o, según lo expresó Lagrange analíticamente, que la expresión

$$dz - p\ dx - q\ dy \tag{72}$$

[40] Para ϕ arbitraria, no es posible, en general, llevar a cabo la eliminación efectiva de a. La integral general queda en el nivel de los conceptos y equivale a una colección de soluciones particulares

[41] *Nouv. Mém. de l'Acad. de Berlin*, 1772 = *Œuvres*, 3, 549-575.

se convierta, multiplicándola por un factor $M(x, y, z)$ apropiado, en una diferencial exacta dN de $N(x, y, z)$. Para ello necesita que se tenga

$$\frac{\partial N}{\partial z} = M, \qquad \frac{\partial N}{\partial x} = -Mp, \qquad \frac{\partial N}{\partial y} = -Mq.$$

Las condiciones de integrabilidad (67) implican para estas ecuaciones que

$$\frac{\partial M}{\partial x} = -\frac{\partial (Mp)}{\partial x}, \qquad \frac{\partial M}{\partial y} = -\frac{\partial (Mq)}{\partial z}, \qquad \frac{\partial (Mp)}{\partial y} = \frac{\partial (Mq)}{\partial x}.$$

Si en la última de estas tres ecuaciones se sustituyen los valores de $\partial M/\partial x$ y $\partial M/\partial y$ dados por las dos primeras, se obtiene

$$\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x} - p\frac{\partial q}{\partial z} + q\frac{\partial p}{\partial z} = 0. \tag{73}$$

Esta es la condición (68) para la integrabilidad de (72), una condición ya conocida, como señala Lagrange. Si en (73) se sustituye q por la función dada Q de x, y, z y p, dicha ecuación se convierte en

$$-Q_p\frac{\partial p}{\partial x} + \frac{\partial p}{\partial y} + (Q - pQ_p)\frac{\partial p}{\partial z} - Q_x - pQ_z = 0. \tag{74}$$

El objetivo de Lagrange es ahora hallar una solución $p = P$ de esta ecuación de primer orden, la cual es *lineal* y cuya solución contiene una constante arbitraria α. Habiendo obtenido esto, integra las dos ecuaciones

$$q - Q(x, y, z, p) = 0, \qquad p - P(x, y, z, \alpha) = 0, \tag{75}$$

que representan $\partial z/\partial x$ y $\partial z/\partial y$ como funciones de x, y, z, y halla una familia de ∞^2 superficies integrales de la ecuación original (70); es decir, determina la solución completa. Hasta este punto, pues, Lagrange ha sustituido el problema de resolver la ecuación no lineal (70) por el de resolver la ecuación *lineal* (74).

Lagrange dio en 1779 su método para resolver las ecuaciones en derivadas parciales de primer orden lineales [42]. Considera la que todavía se denomina ecuación lineal de Lagrange

$$Pp + Qq = R, \tag{76}$$

[42] *Nouv. Mém. de l'Acad de Berlin*, 1779 = *Œuvres*, 4, 585-634.

donde P, Q y R son funciones de x, y, z; la ecuación se dice que es no homogénea debido a la presencia del término R. Esta ecuación está íntimamente ligada a la ecuación homogénea en tres variables

$$P\frac{\partial f}{\partial x} + Q\frac{\partial f}{\partial y} + R\frac{\partial f}{\partial z} = 0. \qquad (77)$$

Lagrange prueba inmediatamente que si $u(x, y, z) = c$ es una solución de (76), entonces $f = u(x, y, z)$ es una solución de (77), y recíprocamente. De aquí que el problema de resolver (76) sea equivalente al de resolver (77). La ecuación (77) está a su vez relacionada con el sistema de ecuaciones diferenciales ordinarias

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R}, \text{ o bien } \frac{dy}{dx} = \frac{Q}{P} \text{ y } \frac{dz}{dx} = \frac{R}{P}. \qquad (78)$$

En efecto, si $f = u(, y, z)$ y $f = v(x, y, z)$ son dos soluciones independientes de (77), entonces $u = c_1$ y $v = c_2$ definen una solución de (78), y recíprocamente. En consecuencia, si podemos determinar las soluciones $u = c_1$, $v = c_2$ de (78), $f = u$ y $f = v$ serán soluciones de (77) y $u = c$ y $v = c$ serán soluciones de (76). Además, se puede demostrar inmediatamente que $f = \phi(u, v)$, donde ϕ es una función arbitraria de u y v, también satisface (77). Así pues, $\phi(u, v) = c$, o, dado que ϕ es arbitraria, $\phi(u, v) = 0$, es solución general de (76). Lagrange desarrolló el esquema anterior en 1779, dando una demostración en 1785 [43]. Merece quizá la pena señalar de paso que Euler sabía que la solución de (77) se puede reducir a la solución de (78).

Si se pone este trabajo sobre ecuaciones lineales en relación con el trabajo de 1772 sobre ecuaciones no lineales, se ve que Lagrange había logrado reducir una ecuación de primer orden arbitraria en x, y y z a un sistema de ecuaciones diferenciales ordinarias simultáneas; no establece el resultado explícitamente, pero se deduce de dicho trabajo. Curiosamente, en 1785 tenía que resolver una ecuación particular de primer orden y afirmó que era imposible hacerlo con los métodos existentes; había olvidado su trabajo de 1772.

A continuación, Paul Charpit (m. 1784) combinó, presumiblemente en 1784, los métodos para ecuaciones no lineales y para ecuaciones lineales a fin de reducir cualquier $f(x, y, z, p, q) = 0$ a un

---

[43] *Nouv. Mém. de l'Acad de Berlin*, 1785 = *Œuvres*, 5, 543-562.

sistema de ecuaciones diferenciales ordinarias. Lacroix afirmó en 1798 que Charpit había presentado en 1784 un artículo (que no fue publicado) en el que reducía las ecuaciones en derivadas parciales de primer orden a sistemas de ecuaciones diferenciales ordinarias. Jacobi encontró sorprendente esa afirmación y expresó su deseo de que el trabajo de Charpit fuese publicado, pero esto nunca se hizo y no sabemos si la afirmación de Lacroix era correcta. Incluso Lagrange pudo haber hecho todo el trabajo y Charpit no haber añadido nada. El método que se expone en los textos modernos, llamado método de Lagrange, de Lagrange-Charpit o de Charpit, consiste en la fusión de las ideas de Lagrange presentadas en los artículos de 1772 y de 1779. Establece que para resolver la ecuación general en derivadas parciales de primer orden $f(x, y, z, p, q)=0$ hay que resolver el sistema de ecuaciones diferenciales ordinarias (las ecuaciones características de $f=0$),

$$\frac{dx}{dt}=f_p, \qquad \frac{dy}{dt}=f_q, \qquad \frac{dz}{dt}=pf_p+qf_q,$$
$$\frac{dp}{dt}=-f_x-f_zp, \qquad \frac{dq}{dt}=-f_y-f_zq. \tag{79}$$

La resolución se efectúa hallando una integral de (79), pongamos $u(x, y, z, p, q)=A$. Entre ésta y $f=0$ se despejan p y q, que se sustituyen en $dz=pdx+qdy$ (ver [72]), integrando esta última por el método utilizado para (65).

El método de Lagrange se denomina a menudo método de las características de Cauchy debido a que la generalización a n variables del método por el que se llega a (79) empleado por Lagrange y Charpit para una ecuación en dos variables independientes presenta algunas dificultades que fueron superadas por Cauchy en 1819 [44].

6. Monge y la teoría de las características

Gaspar Monge (1746-1815) introdujo el lenguaje de la geometría allí donde Lagrange había trabajado de un modo puramente analíti-

---

[44] *Bull. de la Societé Philomathique*, 1819, 10-21; ver también *Exercices d'analyse et de phys. math.*, 2, 238-272 = *Œuvres*, (2), 12, 272-309.

co. Su trabajo en ecuaciones en derivadas parciales no fue tan importante como el de Euler, Lagrange o Legendre, pero inició el movimiento para interpretar geométricamente los trabajos analíticos, introduciendo de ese modo muchas ideas fructíferas. Vio que, así como problemas que involucran curvas conducen a ecuaciones diferenciales ordinarias, así problemas que involucran superficies conducen a ecuaciones en derivadas parciales. De manera más general, geometría y análisis eran para Monge una sola materia, mientras que para los demás matemáticos del siglo eran dos ramas distintas con sólo algunos puntos de contacto. Monge comenzó su trabajo en 1770, pero no publicó hasta mucho más tarde.

Fue primeramente en el campo de las ecuaciones en derivadas parciales de primer orden donde Monge no sólo introdujo la interpretación geométrica sino que puso énfasis en un nuevo concepto, el de curvas características [45]. Sus ideas sobre las características y sobre las integrales en tanto que envolventes no fueron comprendidas por sus contemporáneos, que las tacharon de principio metafísico; la teoría de características, sin embargo, habría de convertirse en una materia importante en trabajos ulteriores. Monge desarrolló sus ideas más completamente en sus lecciones y en subsiguientes publicaciones, especialmente en las *Feuilles d'analyse appliquée à la géometrie* (1795). La mejor manera de ilustrar sus ideas es con su propio ejemplo.

Consideremos la familia biparamétrica de esferas con radio constante R y centro en cualquier punto del plano XY. La ecuación de esta familia es

$$(x-a)^2 + (y-b)^2 + z^2 = R^2. \tag{80}$$

Esta ecuación es la integral completa de la ecuación en derivadas parciales de primer orden no lineal

$$z^2(p^2 + q^2 + 1) = R^2, \tag{81}$$

ya que (80) contiene dos constantes arbitrarias, a y b, y satisface claramente (81). La subfamilia de esferas introducida al hacer $b = \phi(a)$ es una familia de esferas con centro en una curva, la curva $y = \phi(x)$ del plano XY. La envolvente de esta familia uniparamétrica de esferas (una superficie tubular) es también una solución de (81). Esta

[45] *Hist. de l'Acad. des Sci., Paris*, 1784, 85-117 y 118-192, pub. 1787.

solución particular se obtiene eliminando α entre

$$(x-a)^2+(y-\phi(a))^2+z^2=R^2 \qquad (82)$$

y la derivada parcial de (82) con respecto a a, a saber,

$$(x-a)+(y-\phi(a))\phi'(a)=0. \qquad (83)$$

Para cada elección particular de α, las ecuaciones (82) y (83) representan cada una de ellas una superficie particular, y por tanto ambas, consideradas simultáneamente, representan una curva llamada curva característica. Esta curva es también la intersección de dos esferas «consecutivas» de la subfamilia y el conjunto de curvas características genera o cubre la envolvente, es decir, la envolvente toca a cada miembro de la su familia a lo largo de la característica. La integral general es el agregado de superficies (envolventes de familias uniparamétricas), cada una de las cuales está generada por un conjunto de curvas características. La envolvente de todas las soluciones de (80), o sea, la solución singular, obtenida por eliminación de a y b entre (80) y las derivadas parciales de (80) con respecto a a y b, es $z=\pm R$.

Las curvas características aparecen también por otro camino. Consideremos dos subfamilias de esferas cuyas envolventes son tangentes a lo largo de una esfera cualquiera; podemos denominar envolventes consecutivas a tales envolventes. La curva de intersección de esas dos envolventes consecutivas es la misma curva característica sobre la esfera que la obtenida considerando miembros consecutivos de cualquiera de las dos subfamilias. Cualquier esfera puede pertenecer a una infinidad de diferentes subfamilias cuyas envolventes son todas distintas, con lo que habrá distintas curvas características sobre la misma esfera. Todas son circunferencias de círculos máximos situados en planos verticales.

Monge dio la forma analítica de las ecuaciones diferenciales de las curvas características, lo que equivale al hecho de que las ecuaciones (79) determinan las curvas características de (69). (Monge utilizó ecuaciones en diferenciales totales para expresar las ecuaciones de las curvas características.)

Monge introdujo también (en 1784) la noción de cono característico. En cualquier punto (x, y, z) del espacio (fig. 22.4), se puede considerar el plano cuya normal tiene la dirección del vector $(p, q,$

en el espacio, caso en el que la ecuación diferencial concreta es $(1 + q^2)r - 2pqs + (1 + p^2)t = 0$. (Ver también cap. 24, sec. 4). Aunque Monge ya había efectuado algún trabajo sobre la ecuación (87), en el presente artículo (1795) fue capaz de resolverla elegantemente por el método que vamos a esbozar.

Partiendo de los resultados inmediatos

$$dz = p\ dx + q\ dy \tag{88}$$

$$dp = r\ dx + s\ dy \tag{89}$$

$$dq = s\ dx + t\ dy \tag{90}$$

y eliminado r y t de (87), (89) y (90) obtuvo la ecuación

$$s(R\ dy^2 - S\ dx\ dy + T\ dx^2) - (R\ dy\ dp + T\ dx\ dq - V\ dx\ dy) = 0. \tag{91}$$

Argumenta entonces que siempre que sea posible resolver simultáneamente

$$R\ dy^2 - S\ dx\ dy + T\ dx^2 = 0 \tag{92}$$

y

$$R\ dy\ dp + T\ dx\ dq - V\ dx\ dy = 0 \tag{93}$$

se satisfará (91) y lo mismo ocurrirá con (87).

La ecuación (92) es equivalente a dos ecuaciones de primer orden:

$$dy - W_1(x, y, z, p, q)\ dx = 0$$

y

$$dy - W_2(x, y, z, p, q)\ dx = 0. \tag{94}$$

Las ecuaciones (88) y (93), junto con una de las de (94), constituyen un sistema de tres ecuaciones en diferenciales totales en las cinco variables x, y, z, p, q. Cuando se pueden resolver estas tres ecuaciones, es posible encontrar dos soluciones:

$$u_1(x, y, z; p, q) = C_1$$

y

$$u_2(x, y, z, p, q) = C_2,$$

sus lecciones de 1806, por una curva característica se puede hacer pasar infinitas superficies integrales que son tangentes una a otra sobre dicha curva.

7. Monge y las ecuaciones de segundo orden no lineales

Además de las ecuaciones de segundo orden lineales que ya hemos examinado, los matemáticos del siglo XVIII tuvieron ocasión de considerar ecuaciones de segundo orden lineales en dos variables independientes más generales e incluso ecuaciones no lineales. Estudiaron, así, la ecuación lineal

$$A\frac{\partial^2 z}{\partial x^2} + B\frac{\partial^2 z}{\partial x \partial y} + C\frac{\partial^2 z}{\partial y^2} + D\frac{\partial z}{\partial x} + E\frac{\partial z}{\partial y} + Fz + G = 0,$$

donde A, B, ..., G son funciones de x e y. Esta ecuación se escribe usualmente en la forma

$$Ar + Bs + Ct + Dp + Eq + Fz + G = 0, \qquad (85)$$

donde las letras r, s, t, p y q tienen el significado obvio. Laplace demostró en 1773 [47] que al ecuación (85) se puede reducir, mediante un cambio de variables, a la forma

$$s + ap + bq + cz + g = 0, \qquad (86)$$

donde a, b, c y g son funciones únicamente de x e y, supuesto que $B^2 - 4AC \neq 0$. Resolvió después la ecuación en términos de una serie.

En sus *Feuilles d'analyse*, Monge consideró la ecuación no lineal

$$Rr + Ss + Tt = V, \qquad (87)$$

en donde R, S, T y V son funciones de x, y, z, p y q, de modo que la ecuación es lineal sólo en las derivadas segundas r, s y t. Este tipo de ecuación aparece en el trabajo de Lagrange sobre superficies mínimas, esto es, superficies de área mínima limitadas por curvas dadas

[47] *Hist. de l'Acad. des Sci., Paris*, 1773, pub. 1777 = Œuvres, 9, 5-68.

en el espacio, caso en el que la ecuación diferencial concreta es $(1 + q^2)r - 2pqs + (1 + p^2)t = 0$. (Ver también cap. 24, sec. 4). Aunque Monge ya había efectuado algún trabajo sobre la ecuación (87), en el presente artículo (1795) fue capaz de resolverla elegantemente por el método que vamos a esbozar.

Partiendo de los resultados inmediatos

$$dz = p\ dx + q\ dy \tag{88}$$

$$dp = r\ dx + s\ dy \tag{89}$$

$$dq = s\ dx + t\ dy \tag{90}$$

y eliminado r y t de (87), (89) y (90) obtuvo la ecuación

$$s(R\,dy^2 - S\,dx\,dy + T\,dx^2) - (R\,dy\,dp + T\,dx\,dq - V\,dx\,dy) = 0. \tag{91}$$

Argumenta entonces que siempre que sea posible resolver simultáneamente

$$R\ dy^2 - S\ dx\ dy + T\ dx^2 = 0 \tag{92}$$

y

$$R\ dy\ dp + T\ dx\ dq - V\ dx\ dy = 0 \tag{93}$$

se satisfará (91) y lo mismo ocurrirá con (87).

La ecuación (92) es equivalente a dos ecuaciones de primer orden:

$$dy - W_1(x, y, z, p, q)\ dx = 0$$

y (94)

$$dy - W_2(x, y, z, p, q)\ dx = 0.$$

Las ecuaciones (88) y (93), junto con una de las de (94), constituyen un sistema de tres ecuaciones en diferenciales totales en las cinco variables x, y, z, p, q. Cuando se pueden resolver estas tres ecuaciones, es posible encontrar dos soluciones:

$$u_1(x, y, z; p, q) = C_1$$

y

$$u_2(x, y, z, p, q) = C_2,$$

y entonces

$$u_1 = \phi\ (u_2), \tag{95}$$

con ϕ arbitraria, es una ecuación en derivadas parciales de primer orden. La ecuación (95) se denomina integral intermedia y su solución general es la solución de (87). Si se puede utilizar la otra ecuación de (94) junto con (88) y (93), se obtiene otra función:

$$u_3 = \phi\ (u_4). \tag{96}$$

En este caso, (95) y (96) se pueden resolver simultáneamente en p y q y estos valores se sustituyen en (88), pudiéndose entonces resolver esta ecuación en diferenciales totales. Este es al menos el esquema general, aunque hay detalles en los que no entraremos. La integración por Monge de la ecuación de las superficies mínimas fue una de sus reivindicaciones de gloria.

Monge introdujo también la teoría de características para la ecuación (87). La ecuación en diferenciales totales de las características es (92), es decir,

$$R\ dy^2 - S\ dx\ dy + T\ dx^2 = 0.$$

Esta ecuación, que aparece en su obra ya en 1784 [48], define en cada punto de una superficie integral dos direcciones que son las direcciones características en ese punto. Por cada punto de una superficie integral pasan dos curvas características, a lo largo de cada una de las cuales se tocan dos superficies integrales consecutivas.

8. Sistemas de ecuaciones en derivadas parciales de primer orden

Los sistemas de ecuaciones en derivadas parciales surgieron por primera vez en los trabajos del siglo XVIII sobre dinámica de fluidos o hidrodinámica. El interés por los fluidos incompresibles —agua, por ejemplo—, estaba motivado por problemas prácticos tales como el diseño de cascos de buques para reducir la resistencia al movi-

[48] *Hist. de l'Acad. des Sci., Paris*, 1784, 118-192, pub. 1787.

miento en el agua y el cálculo de las mareas, del flujo de los ríos, del flujo del agua en surtidores y de la presión del agua sobre los costados de un buque. Los trabajos sobre fluidos compresibles, el aire en particular, intentaban analizar la acción del aire sobre las velas de un navío, el diseño de las aspas de los molinos de viento y la propagación del sonido. Los trabajos que examinamos anteriormente sobre la propagación del sonido supusieron, desde el punto de vista histórico, una aplicación de los trabajos sobre hidrodinámica particularizados a ondas de pequeña amplitud.

Después de haber tratado en 1752 los fluidos incompresibles en un artículo titulado «Principios del movimiento de fluidos» [49], Euler generalizó este trabajo en un artículo de 1755 titulado «Principios generales del movimiento de fluidos» [50]. Formuló en él las todavía célebres ecuaciones del movimiento de un fluido perfecto (no viscoso), compresible e incompresible. El fluido se contempla como un continuo y las partículas como puntos matemáticos. Euler considera la fuerza que actúa sobre un pequeño volumen del fluido de densidad ρ, sujeto a una presión p, y una fuerza externa de componente P, Q y R por unidad de masa.

En uno de los enfoques de la dinámica de fluidos que Euler creó, conocido en la literatura como descripción espacial, las componentes u, v y w de la velocidad del fluido en cada punto de éste están dadas por

$$u = u(x, y, z, t), \qquad v = v(x, y, z, t), \qquad w = w(x, y, z, t). \tag{97}$$

Se tendrá

$$du = \frac{\partial u}{\partial x}\, dx + \frac{\partial u}{\partial y}\, dy + \frac{\partial u}{\partial z}\, dz + \frac{\partial u}{\partial t}\, dt.$$

En el tiempo dt, la partícula que está en (x, y, z) se desplazará una distancia $u\,dt$ en la dirección del eje x, $v\,dt$ en la del eje y, y $w\,dt$ en la del eje z. En consecuencia, los cambios dx, dy, dx producidos en la expresión de du estarán dados por esas cantidades, y se tendrá:

$$du = \frac{\partial u}{\partial x}\, u\, dt + \frac{\partial u}{\partial y}\, v\, dt + \frac{\partial u}{\partial z}\, w\, dt + \frac{\partial u}{\partial t}\, dt,$$

[49] *Novi Comm. Acad. Sci. Petrop.*, 6, 1756/57, 271-311, pub. 1761 = *Opera*, (2), 12, 133-168.

[50] *Hist. de l'Acad de Berlin*, 11, 1755, 274-315, pub. 1757 = Opera, (2), 12, 54-91.

o bien

$$\frac{du}{dt} = u\,\frac{\partial u}{\partial x} + v\,\frac{\partial u}{\partial y} + w\,\frac{\partial u}{\partial z} + \frac{\partial u}{\partial t}, \qquad (98)$$

teniéndose las correspondientes expresiones para dv/dt y dw/dt. Estas cantidades dan lo que hoy se llama razón de cambio convectiva de la velocidad en (x, y, z) o aceleración convectiva. Calculando las fuerzas que actúan sobre la partícula situada en (x, y, z) y aplicando la segunda ley de Newton, Euler obtiene el sistema de ecuaciones diferenciales:

$$P - \frac{1}{\varrho}\,\frac{\partial p}{\partial x} = \frac{du}{dt}$$

$$Q - \frac{1}{\varrho}\,\frac{\partial p}{\partial y} = \frac{dv}{dt} \qquad (99)$$

$$R - \frac{1}{\varrho}\,\frac{\partial p}{\partial z} = \frac{dw}{dt}.$$

Euler también generalizó la ecuación diferencial de continuidad de D'Alembert (45) y obtuvo para un fluido compresible la ecuación

$$\frac{\partial \varrho}{\partial t} + \frac{\partial(\varrho u)}{\partial x} + \frac{\partial(\varrho v)}{\partial y} + \frac{\partial(\varrho w)}{\partial z} = 0. \qquad (100)$$

Hay cuatro ecuaciones y cinco incógnitas pero falta por especificar la presión como función de la densidad, o sea, la ecuación de estado.

En el artículo de 1755, Euler afirma: «Y si no nos está permitido llegar a un completo conocimiento en lo que concierne al movimiento de los fluidos, no es a la mecánica o a la insuficiencia de los principios del movimiento conocidos a los que hemos de atribuir la causa. Es el análisis el que nos abandona aquí, pues toda la teoría del movimiento de fluidos ha sido reducida precisamente a la resolución de fórmulas analíticas.» Desgraciadamente, el análisis era todavía muy débil para hacer gran cosa con estas ecuaciones y Euler procedió a discutir algunas soluciones especiales. También escribió otros artículos sobre el tema y trató de la resistencia que encuentran los buques y de la propulsión de éstos. Las ecuaciones de Euler no

son las ecuaciones definitivas de la hirodinámica; no tienen en cuenta la viscosidad, que fue introducida setenta años más tarde por Navier y Stokes (cap. 28, sec. 7).

También Lagrange trabajó sobre el movimiento de fluidos. En la primera edición de su *Mécanique analytique*, que contiene parte del trabajo, dio las ecuaciones fundamentales de Euler y las generalizó. Atribuye en esa obra el mérito a D'Alembert y no da ninguno a Euler. Afirma también que las ecuaciones del movimiento de un fluido son demasiado difíciles para ser tratadas por el análisis y que sólo los casos de movimientos infinitamente pequeños son susceptibles de cálculos rigurosos.

Las ecuaciones de la hidrodinámica fueron, en el siglo XVIII, la principal inspiración para la investigación matemática en el campo de los sistemas de ecuaciones en derivadas parciales. En realidad, se avanzó poco en ese siglo en la resolución de sistemas.

9. El desarrollo de las ecuaciones en derivadas parciales como disciplina matemática

Hasta 1765, las ecuaciones en derivadas parciales aparecían únicamente en la resolución de problemas físicos. El primer artículo dedicado a un trabajo puramente matemático sobre tales ecuaciones se debe a Euler: «Recherches sur l'intégration de l'equation $\left(\frac{d\,dz}{dt^2}\right) = aa\left(\frac{d\,dz}{dx^2}\right) + \frac{b}{x}\left(\frac{dz}{dx}\right)$.» [51] Poco después, Euler publicó un tratado sobre la materia en el tercer volumen de sus *Institutiones Calculi Integralis* [52].

Antes del trabajo de D'Alembert de 1747 sobre la cuerda vibrante, las ecuaciones en derivadas parciales eran conocidas como ecuaciones de estado y sólo se buscaban soluciones especiales. Después de ese trabajo y del libro de D'Alembert sobre las causas generales de los vientos (1746), los matemáticos comprendieron la diferencia entre soluciones especiales y la solución general. Sin embargo, una vez conscientes de esa distinción, creyeron, aparentemente, que debían ser más importantes las soluciones generales. En el primer vo-

[51] *Misc. Taur.*, 3_2 1762/65, 60-91, pub. 1766 = *Opera*, (1), 23, 47-73.
[52] 1770 = *Opera*, (1), 13.

lumen de su *Mécanique céleste* (1799), Laplace todavía se quejaba de que no se pudiese integrar de manera general la ecuación del potencial en coordenadas esféricas. No se llegó a valorar en este siglo que soluciones generales como las obtenidas por Euler y D'Alembert en el problema de la cuerda vibrante no eran tan útiles como las soluciones particulares que satisfacen determinadas condiciones iniciales y de contorno.

Sí se dieron cuenta los matemáticos de que las ecuaciones en derivadas parciales no implicaban nuevas técnicas operativas, pero que diferían de las ecuaciones diferenciales ordinarias en que pueden aparecer en su resolución funciones arbitrarias. Esperaban determinar éstas reduciendo las ecuaciones en derivadas parciales a ordinarias. Laplace (1773) y Lagrange (1784) afirmaban claramente que consideraban integrada una ecuación en derivadas parciales cuando quedaba reducida a un problema de ecuaciones diferenciales ordinarias. Una alternativa, como la utilizada por Daniel Bernoulli para la ecuación de ondas y Laplace para la ecuación del potencial, consistía en buscar desarrollos en serie de funciones especiales.

El principal logro de los trabajos del siglo XVIII sobre ecuaciones en derivadas parciales fue poner de manifiesto su importancia en los problemas de elasticidad, hidrodinámica y atracción gravitatoria. Excepto en el caso del trabajo de Lagrange sobre las ecuaciones de primer orden, no se desarrollaron métodos generales ni se supieron valorar las potencialidades del método de desarrollos en serie de funciones especiales. Los esfuerzos se dirigían a resolver ecuaciones especiales que aparecían en problemas físicos. Quedaba por elaborar la teoría de la resolución de ecuaciones en derivadas parciales, y la materia en su conjunto estaba todavía en su infancia.

Bibliografía

Burkhardt, H.: «Entwicklungen nach oscillirenden Funktionen und Integration der Differentialgleichungen der mathematischen Physik», *Jahres. der Deut. Math.-Verein.*, 10, 1908, 1-1804.

Burkhardt, H. y W. Franz Meyer: «Potentialtheorie», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, 2, A7b, 464-503.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teub-

ner, 1898 y 1924; Johnson Reprint Corp., vol. 3, 858-878, vol. 4, 873-1047.

Euler, Leonhard: *Opera Omnia*, Orell Füssli, (1), vols. 13 (1914) y 23 (1938); (2), vols. 10, 11, 12 y 13 (1947-55).

Lagrange, Joseph-Louis: *Œuvres*, Gauthier-Villars, 1868-70, los artículos relevantes en vols. 1, 3, 4 y 5.

Laplace, Pierre-Simon: *Œuvres complètes*, Gauthier-Villars, 1893-94, los artículos relevantes en vols. 9 y 10.

Montucla, J. F.: *Histoire des mathématiques* (1802), Albert Blanchard (reimpresión), 1960, vol. 3, págs. 342-352.

Langer, Rudolph E.: «Fourier Series: The Genesis and Evolution of a Theory», *Amer. Math. Monthly*, 54, núm. 7, parte 2, 1947.

Taton, René: *L'Œuvre scientifique de Monge*, Presses Universitaires de France, 1951.

Todhunter, Isaac: *A History of the Mathematical Theories of Attraction and the Figure of the Earth* (1873), Dover (reimpresión), 1962.

Truesdell, Clifford E.: *Introduction to Leonhardi Euleri Opera Omnia, vol. X et XI Seriei Secundae*, en Euler, *Opera Omnia*, (2), 11, parte 2, Orell Füssli, 1960.

—: *Editor's Introduction* en Euler, *Opera Omnia*, (2), vol. 12, Orell Füssli, 1954.

—: *Editor's Introduction* en Euler, *Opera Omnia*, (2), vol. 13, Orell Füssli, 1956.

Capítulo 23

GEOMETRIA ANALITICA Y DIFERENCIAL EN EL SIGLO XVIII

Aparentemente, en ocasiones, la geometría puede tomar ventaja sobre el análisis, pero de hecho lo precede únicamente como un sirviente cuando adelanta a su amo para limpiar el recorrido e iluminarlo en su camino.

JAMES JOSEPH SYLVESTER

1. Introducción

La exploración de problemas físicos lleva inevitablemente a la búsqueda de un mejor conocimiento de las curvas y superficies, ya que las trayectorias de objetos en movimiento son curvas y los objetos en sí mismos son objetos de tres dimensiones limitados por superficies. Los matemáticos, ya de por sí entusiasmados por el método de la geometría de coordenadas y por la fuerza del cálculo, abordaron los problemas geométricos con estas dos poderosas herramientas. Los resultados impresionantes del siglo fueron obtenidos en el área ya establecida de la geometría de coordenadas y el nuevo campo creado por la aplicación del cálculo a los problemas geométricos, que se llamó geometría diferencial.

2. Geometría analítica básica

La geometría analítica de dos dimensiones fue explorada extensamente durante el siglo XVIII. Los avances en la geometría analítica plana elemental ya han sido resumidos. Mientras que Newton y

Jacques Bernoulli habían introducido y empleado lo que son esencialmente las coordenadas polares para curvas especiales (cap. 15, sec. 5), en 1729, Jacob Hermann no solamente proclamó su utilidad general, sino que las aplicó libremente al estudio de las curvas, proporcionando a su vez la transformación de coordenadas polares a rectangulares. Estrictamente hablando, Hermann usó variables, ϱ, cos θ, sen θ, que él designó por z, n y m. Euler extendió el uso de las coordenadas polares y utilizó explícitamente la notación trigonométrica; con él, prácticamente, el sistema es el moderno.

A pesar de que algunos autores del siglo XVII —por ejemplo, Jan de Witt (1625-72) en su *Elementa Curvarum Linearum* (1659)— redujeron algunas ecuaciones de segundo grado en x e y a formas canónicas, James Stirling, en su *Lineae Tertii Ordinis Neutonianae* (1717), redujo la ecuación general de segundo grado en x e y a las diversas formas canónicas.

En su *Introductio* (1748), Euler incorpora la representación paramétrica de las curvas, donde x e y son coordenadas, en términos de una tercera variable. En este famoso texto, Euler trató sistemáticamente la geometría plana con coordenadas.

Las sugerencias de una geometría de coordenadas de tres dimensiones pueden ser encontradas, como sabemos, en las obras de Fermat, Descartes y La Hire. El desarrollo efectivo fue trabajo del siglo XVIII. Aunque alguna parte de los trabajos iniciales, por ejemplo, el de Pilot y Clairaut, están ligados al desarrollo de la geometría diferencial, consideraremos por ahora únicamente la geometría de coordenadas.

La primera tarea consistió en mejorar la sugerencia de La Hire respecto a un sistema coordenado de tres dimensiones. Jean Bernoulli, en una carta a Leibniz de 1715, introduce los planos de tres coordenadas que utilizamos hoy en día. A través de aportaciones demasiado detalladas para incluirlas aquí, Antonie Parent (1666-1716), Jean Bernoulli, Clairaut y Jacob Hermann clarificaron la noción de que una superficie puede ser representada por una ecuación en tres coordenadas. Clairaut, en su libro *Recherche sur les courbes a double courbure* (*Investigaciones sobre las curvas de doble curvatura*, 1731) no sólo proporcionó las ecuaciones de algunas superficies, sino que mostró que dos de esas ecuaciones son necesarias para describir una curva en el espacio. También vio que ciertas combinaciones de las ecuaciones de dos superficies que pasan a través de una curva —por ejemplo, la suma—, proporcionan la ecuación de

otra superficie que contiene la curva. Usando este hecho, explica cómo se pueden obtener las ecuaciones de las proyecciones de estas curvas, o, equivalentemente, las ecuaciones de los cilindros perpendiculares a los planos de proyección.

Las superficies cuadráticas, por ejemplo la esfera, el cilindro, la paraboloide, el hiperboloide de dos hojas y el elipsoide, por supuesto eran conocidas geométricamente antes de 1700; de hecho, algunas de ellas aparecen en las obras de Arquímedes. En su libro de 1731, Clairaut dio las ecuaciones de algunas de estas superficies. También mostró que una ecuación homogénea en x, y y z (donde todos los términos son del mismo grado) representa un cono con vértice en el origen. A este resultado, Jacob Hermann, en un ensayo de 1732 [1], añadió que la ecuación $x^2 + y^2 = f(z)$ es una superficie de revolución alrededor del eje z. Tanto Clairaut como Hermann estaban primordialmente interesados en la forma de la Tierra, que en su tiempo se pensaba tenía cierta conformación de elipsoide.

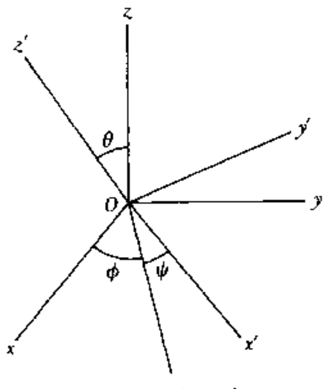
A pesar de que Euler había realizado un trabajo previo sobre las ecuaciones de superficies, es en el capítulo cinco de su apéndice al segundo volumen de su *Introductio* [2] donde aborda sistemáticamente la geometría plana tridimensional. Presenta mucho de lo que ya se había realizado hasta entonces y, porteriormente, estudia la ecuación general de segundo grado en tres variables

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + kz = l. \tag{1}$$

Ahora, Euler busca cambiar los ejes para reducir esta ecuación a las formas que resultan de tomar los ejes principales de las superficies cuadráticas representadas por (1) como ejes coordenados. Introduce la transformación del sistema xyz al sistema $x'y'z'$, cuyas ecuaciones están expresadas (fig. 23.1) en términos de los ángulos ϕ, ψ y θ. El ángulo ϕ es medido en el plano xy a partir del eje x a la línea de nodos, que es la línea en que el plano $x'y'$ corta el plano xy. El ángulo ψ se mide en el plano $x'y'$ y localiza a x' con respecto a las líneas de los nodos. El ángulo que se muestra es θ, por lo que las ecuaciones de transformación, incluyendo traslación, son:

[1] *Comm. Acad. Sci. Petrop.*, 6, 1732/33, 36-67, pub. 1738.

[2] *Opera*, (1), 9.



Línea de nodos.

FIGURA 23.1.

$$\begin{aligned}
x &= x_0 + x'(\cos\psi\cos\phi - \cos\theta\,\text{sen}\,\psi\,\text{sen}\,\phi) \\
&\qquad - y'(\cos\psi\,\text{sen}\,\phi + \cos\theta\,\text{sen}\,\psi\,\text{sen}\,\phi) + z'\,\text{sen}\,\theta\,\text{sen}\,\phi \\
y &= y_0 + x'(\text{sen}\,\psi\cos\phi + \cos\theta\cos\psi\,\text{sen}\,\phi) \qquad (2) \\
&\qquad - y'(\text{sen}\,\psi\,\text{sen}\,\phi - \cos\theta\cos\psi\,\text{sen}\,\phi) - z'\,\text{sen}\,\theta\,\text{sen}\,\phi \\
z &= z_0 + x'\,\text{sen}\,\theta\,\text{sen}\,\phi + y'\,\text{sen}\,\theta\cos\phi + z'\cos\theta.
\end{aligned}$$

Euler usa esta transformación para reducir (1) a formas canónicas y obtiene seis casos distintos: cono, cilindro, elipsoide, hiperboloide de una y dos hojas, paraboloide hiperbólico (que él descubrió) y cilindro parabólico. Como Descartes, Euler mantuvo que la clasificación por medio del grado de la ecuación era el principio correcto; la razón de Euler era que el grado es invariante bajo las transformaciones lineales.

Después de continuar trabajando en este problema de cambio de ejes, escribió otro ensayo [3], en el cual considera la transformación que llevará de $x^2 + y^2 + z^2$ a $x'^2 + y'^2 + z'^2$. El aquí —y Lagrange un poco más tarde, en un ensayo sobre la atracción de los esferoides [4]— proporcionó la forma simétrica para la rotación de ejes, la transformación lineal ortogonal homogénea

$$\begin{aligned}
x &= \lambda x' + \mu y' + \nu z' \\
y &= \lambda' x' + \mu' y' + \nu' z' \\
z &= \lambda'' x' + \mu'' y' + \nu'' z',
\end{aligned}$$

[3] *Novi Comm. Acad. Sci. Petrop.*, 15, 1770, 75-106, pub. 1771 = *Opera*, (1), 6, 287-315.

[4] *Nouv. Mém. de l'Acad de Berlin*, 1773, 85-120 = *Œuvres*, 3, 619-658.

donde

$$\begin{array}{ll} \lambda^2 + \lambda'^2 + \lambda''^2 = 1 & \lambda\mu + \lambda'\mu' + \lambda''\mu'' = 0 \\ \mu^2 + \mu'^2 + \mu''^2 = 1 & \lambda\nu + \lambda'\nu' + \lambda''\nu'' = 0 \\ \nu^2 + \nu'^2 + \nu''^2 = 1 & \mu\nu + \mu'\nu' + \mu''\nu'' = 0. \end{array}$$

Las λ, μ y ν, primas y no primas, son los cosenos directores en la terminología actual.

Los escritos de Gaspar Monge contienen una gran cantidad de geometría analítica tridimensional. Su contribución sobresaliente a la geometría analítica como tal se encuentra en un ensayo de 1802, escrito junto con su discípulo Jean Nicolas Pierre Hachette (1769-1834), «Application de l'algèbre a la géométrie» [5] («Aplicación del álgebra a la geometría»). Los autores muestran que cada sección plana de una superficie de segundo grado es una curva de segundo grado, y que planos paralelos cortan en curvas similares análogamente situadas. Estos resultados son análogos a los teoremas geométricos de Arquímedes. Los autores también muestran que el hiperboloide de una hoja y el paraboloide hiperbólico son superficies regladas, esto es, cada una puede ser generada de dos maneras diferentes por el movimiento de una línea, o bien cada superficie está formada por dos sistemas de líneas. El resultado en el hiperboloide de una hoja era conocido en 1669 por Christopher Green, quien afirmaba que esta figura podía ser generada girando una línea alrededor de otra que no se encontrara en el mismo plano. Con el trabajo de Euler, Lagrange y Monge, la geometría analítica se convirtió en una rama de las matemáticas independiente y acabada.

3. Curvas planas de grado superior

La geometría analítica descrita hasta ahora fue dedicada a las curvas y superficies de primero y segundo grado. Por lo tanto, era natural investigar las curvas de ecuaciones de grados superiores. De hecho, Descartes había discutido con anterioridad tales ecuaciones y sus curvas. El estudio de curvas de grados mayores que dos llegó a ser conocido como la teoría de curvas planas de grado superior, a pesar de ser parte de la geometría analítica. Las curvas estudiadas en

[5] *Jour. de l'Ecole Poly.*, cahier 1802, 143-169.

el siglo XVIII eran algebraicas; es decir, sus ecuaciones están dadas por $f(x, y) = 0$, donde f es un polinomio en x e y. El grado u orden es el grado máximo de los términos.

El primer estudio amplio de las curvas planas superiores lo realizó Newton. Impresionado por el plan de Descartes en cuanto a clasificar curvas de acuerdo al grado de sus ecuaciones y entonces estudiar sistemáticamente las de cada grado por métodos adaptados a ese grado, Newton emprendió el estudio de curvas de tercer grado. Este trabajo apareció en su *Enumeratio Linearum Tertii Ordinis (Enumeración de líneas de tercer grado)*, que fue publicado en 1704 como un apéndice a la edición inglesa de su *Opticks (Optica)*, pero ya había sido compuesto en 1676. Aunque el uso de valores negativos de x e y aparece en el trabajo de La Hire y Wallis, Newton no emplea únicamente los dos ejes y valores negativos de x e y, sino que traza los cuatro cuadrantes.

Newton mostró cómo todas las curvas comprendidas por la ecuación general de tercer grado

$$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + jy + k = 0 \quad (3)$$

pueden, con un cambio de ejes, ser reducidas a una de las cuatro formas siguientes:

$$\begin{array}{ll} \text{(a)} & xy^2 + ey = ax^3 + bx^2 + cx + d \\ \text{(b)} & xy = ax^3 + bx^2 + cx + d \\ \text{(c)} & y^2 = ax^3 + bx^2 + cx + d \\ \text{(d)} & y = ax^3 + bx^2 + cx + d. \end{array}$$

La tercera clase, que Newton llamó parábolas divergentes, contiene cinco especies de curvas, cuyos tipos se muestran en la figura 23.2. Las especies se distinguen por la naturaleza de las raíces de la cúbica en el miembro derecho, como sigue: todas reales y distintas; dos raíces complejas; todas reales, pero dos iguales y la doble raíz mayor o menor que la raíz simple; y las tres iguales. Newton afirmaba que cada curva cúbica puede ser obtenida mediante la proyección de uno de estos cinco tipos de curvas a partir de un punto y por una sección de la proyección.

Newton no proporcionó las pruebas de muchas sus afirmaciones en su *Enumeratio*. En su *Lineae*, James Stirling demostró o refutó de otras maneras la mayoría de las afirmaciones pero no el teorema

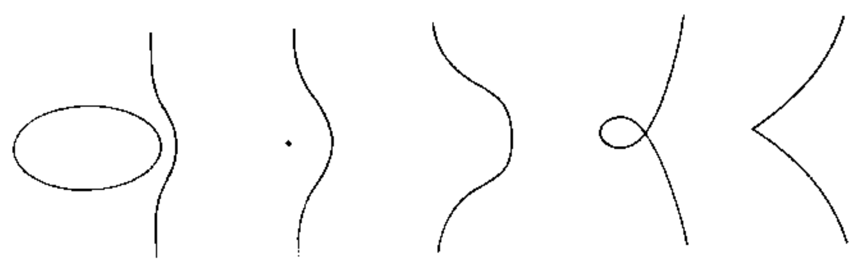


FIGURA 23.2

de la proyección, que Clairaut [6] y François Nicole demostraron (1683-1758) [7]. También, mientras que Newton reconocía setenta y dos especies de curvas de tercer grado, Stirling añadió cuatro más y el abad Jean-Paul de Gua de Malves, en un pequeño libro de 1740 titulado *Usage de l'analyse de Descartes pour decouvrir sans le secours du calcul differential... (Uso del análisis de Descartes para descubrir...)*, añadió dos más.

El trabajo de Newton sobre curvas de tercer grado estimuló mucho la labor de otros sobre curvas planas superiores. El tema de la clasificación de curvas de tercer y cuarto grado de acuerdo con uno u otro principio continuó interesando a matemáticos de los siglos XVIII y XIX. El número de clases encontradas variaba con los métodos de clasificación.

Como es evidente a partir de las figuras de las cinco especies de curvas cúbicas de Newton, las curvas de ecuaciones de grado superior exhiben muchas peculiaridades no encontradas en curvas de primero y segundo grado. Las peculiaridades elementales, llamadas puntos singulares, son puntos de inflexión y puntos múltiples. Antes de seguir, veamos cómo son algunos de éstos.

Los puntos de inflexión son familiares en el cálculo. Un punto en el cual existen dos o más tangentes que pueden coincidir es llamado punto múltiple. En tal punto dos o más ramas de la curva se intersecan: si dos ramas se intersecan en un punto múltiple, se llama punto doble; si tres ramas se intersecan, entonces se llama un punto triple, y así en adelante.

Si tomamos la ecuación de una curva algebraica

$$f(x, y) = 0,$$

[6] *Mém. de l'Acad. des Sci., Paris*, 1731, 490-493, pub. 1733.
[7] *Mém. de l'Acad. des Sci., Paris*, 1731, 494-510, pub. 1733.

siendo f un polinomio en x e y, podemos mediante una traslación suprimir siempre el término constante. Si se hace esto y si hay términos de primer grado en f, digamos $a_1x_1 + b_1y$, entonces $a_1x + b_1y = 0$ proporciona la ecuación de la tangente a la curva en el origen. El origen no es un punto múltiple en este caso. Si no existen términos de primer grado, y si $a_2x^2 + b_2xy + c_2y^2$ son términos de segundo grado, entonces surgen varios casos. La ecuación $a_2x^2 + b_2xy + c_2y^2 = 0$ puede representar dos rectas diferentes. Estas rectas son tangentes en el origen (esto puede ser probado), y ya que hay dos tangentes diferentes el origen es un punto doble, llamado nodo. La ecuación de la lemniscata (fig. 23.3) es

$$a^2(y^2 - x^2) + (y^2 + x^2)^2 = 0 \tag{4}$$

y los términos de segundo grado dan $y^2 - x^2 = 0$. Entonces, $y = x$ e $y = -x$ son las ecuaciones de las tangentes. Del mismo modo, el folio de Descartes (fig. 23.4) tiene la ecuación

$$x^3 + y^3 = 3\ axy \tag{5}$$

y las tangentes en el origen, que es un nodo, están dadas por $x = 0$ e $y = 0$.

Cuando las dos rectas tangentes coinciden, la única recta es considerada como una tangente doble y las dos ramas de la curva se tocan una a otra en el punto de tangencia, el cual es llamado cúspide. (Algunas veces las cúspides están incluidas entre los puntos dobles.) Así la parábola semicúbica (fig. 23.5)

$$ay^2 = x^3 \tag{6}$$

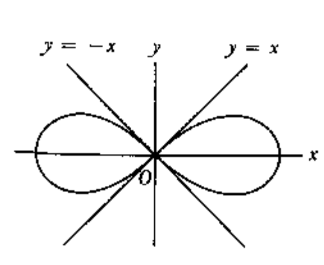


FIGURA 23.3. *Lemniscata.*

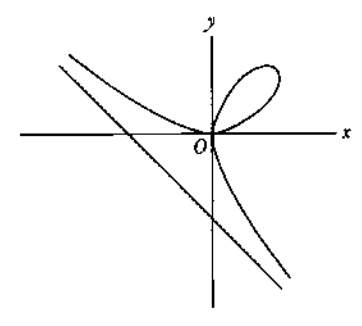


FIGURA 23.3. *Folio de Descartes.*

tiene una cúspide en el origen, y la ecuación de las dos tangentes coincidentes es $y^2 = 0$. Sobre la curva $(y - x^2)^2 = x^5$ (fig. 23.6) el origen es una cúspide. Aquí ambas ramas están del mismo lado de la tangente doble, la cual es $y = 0$. De Gua, en su *Usage*, había intentado mostrar que este tipo de cúspide pudiera no aparecer, pero Euler [8] proporcionó muchos ejemplos. Una cúspide es llamada también punto estacionario o punto de retroceso, porque un punto moviéndose a lo largo de una curva debe detenerse en una cúspide antes de continuar su movimiento.

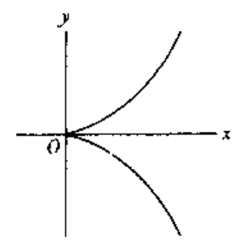


FIGURA 23.5. *Parábola semicúbica.*

FIGURA 23.6

Cuando las dos tangentes son imaginarias, el punto doble es llamado punto conjugado. Las coordenadas del punto satisfacen la ecuación de la curva, pero el punto está aislado del resto de la curva. De aquí que la curva (fig. 23.7) $y^2 = x^2\ (2x - 1)$ tenga un punto conjugado en el origen. La ecuación de la tangente doble es $y^2 = -x^2$ y las tangentes son imaginarias.

La curva $ay^3 - 3ax^2y = x^4$ (fig. 23.8) tiene un punto triple en el origen. La ecuación de las tres tangentes es

$$ay^3 - 3ax^2y = 0$$

o bien $y = 0$ e $y = \pm x\ \sqrt{3}$.

La curva $ay^4 - ax^2\ y^2 = x^5$ (fig. 23.9) tiene un punto cuádruple en el origen. Este origen es una combinación de un nodo y una cúspide. Las tangentes son $y = 0$, $y = 0$ e $y = \pm x$.

Las curvas de tercer grado (orden) pueden tener un punto doble

---

[8] *Mém. de l'Acad de Berlín*, 5, 1749, 203-221, pub. 1751 = *Opera*, (1), 27, 236-252.

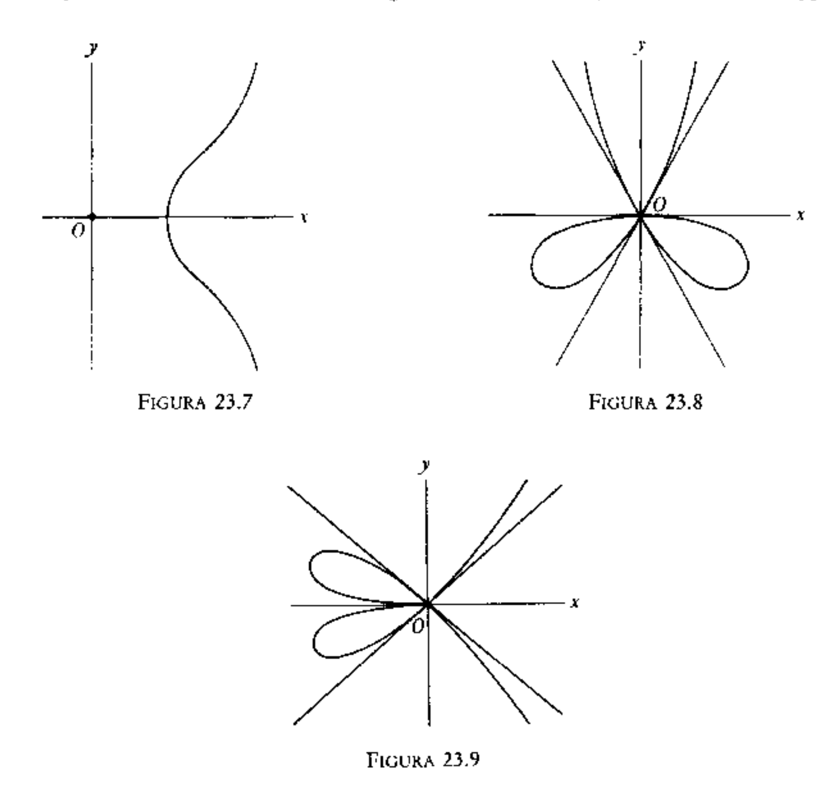


FIGURA 23.7

FIGURA 23.8

FIGURA 23.9

(el cual puede ser una cúspide) pero ningún otro punto múltiple. Existen, desde luego, cúbicas sin puntos dobles.

Para retomar la historia propiamente dicha, muchos de estos puntos especiales o singulares de las curvas fueron estudiados por Leibniz y sus sucesores. Las condiciones analíticas para dichos puntos, tales como que $\ddot{y}=0$ en un punto de inflexión y que $\dot{y}$ sea indeterminada en un punto doble, eran conocidas incluso por los fundadores del cálculo.

Clairaut, en su libro de 1731, al que nos habíamos referido, supone que una curva de tercer grado no puede tener más de tres puntos de inflexión reales y que al menos debe tener uno. De Gua, en el *Usage*, probó que si una curva de tercer grado tiene tres puntos de inflexión reales, una recta que une dos de ellos pasa por el tercero. Este teorema se le atribuye comúnmente a Maclaurin. De Gua también investigó los puntos dobles y proporcionó la condición de

que si $f(x, y) = 0$ es la ecuación de la curva, entonces f_x y f_y deben ser 0 en el punto doble. Un punto de k-pliegue está caracterizado por la anulación de todas las derivadas parciales hasta las de $(k - 1)$-ésimo orden. De Gua mostró que las singularidades están compuestas por cúspides, puntos ordinarios y puntos de inflexión. Además, trató los puntos medios de las curvas, la forma de ramas extendiéndose al infinito y las propiedades de tales ramas.

En su *Geometria Organica* (1720), escrita cuando sólo tenía diecinueve años, Maclaurin probó que el número máximo de puntos dobles de una curva irreducible de n-ésimo grado es $(n - 1)(n - 2)/2$. Para este propósito contó un punto de k-pliegue como $k(k - 1)/2$ puntos dobles, y proporcionó también cotas superiores para el número de puntos de mayor multiplicidad de cada clase. Más adelante introdujo la noción de deficiencia (posteriormente llamada género) de una curva algebraica como el máximo número posible de puntos dobles menos el número verdadero. Entre las curvas, aquellas cuya deficiencia es 0, o poseedoras del máximo número posible de puntos dobles, recibieron gran atención. Tales curvas también son llamadas racionales o unicursales. Geométricamente, una curva unicursal puede ser recorrida por el movimiento continuo de un punto móvil (el cual puede, sin embargo, pasar a través del punto del infinito). De aquí que las cónicas, incluyendo la hipérbola, sean curvas unicursales.

En su *Method of Fluxions (Método de fluxiones)*, Newton dio un método, al que uno se refiere comúnmente como el diagrama de Newton o el paralelogramo de Newton, para determinar representaciones en series de las diversas ramas de una curva en un punto de inflexión (cap. 20, sec. 2). En el *Usage*, De Gua reemplazó el paralelogramo de Newton por un *triangle algébrique* (triángulo algebraico). Entonces, si el origen es un punto singular, para una x pequeña la ecuación de una curva algebraica se descompone en factores de la forma $y^m - Ax^n$, donde m es un entero positivo y n un entero. Las ramas de la curva están dadas por los factores para los cuales n también es positiva. Euler notó (1749) que De Gua había ignorado ramas imaginarias.

Gabriel Cramer (1704-52), en su *Introduction a l'analyse des lignes courbes algébriques* (*Introducción al análisis de las líneas curvas algebraicas*, 1750), también abordó el desarrollo de y en términos de x cuando y y x están dadas por una función implícita, esto es, que $f(x, y) = 0$, determina la expresión en serie para cada rama de la curva, particularmente las ramas que se extienden al infinito. Cramer

trató y como una serie ascendente y descendente de las potencias de x. Al igual que De Gua, usó el triángulo en lugar del paralelogramo de Newton, y como otros, también ignoró las ramas imaginarias de las curvas.

La conclusión resultante del trabajo de obtener expansiones en serie para cada rama de la curva emanando de un punto múltiple fue obtenida mucho más tarde por Victor Puiseux (1820-83) [9] y se conoce como el teorema de Puiseux: el entorno total de un punto (x_0, y_0) de una curva algebraica plana puede ser expresado por un número finito de desarrollos

$$y - y_0 = a_1(x - x_0)^{q_1/q_0} + a_2(x - x_0)^{q_2/q_0} + \cdots \tag{7}$$

Estos desarrollos convergen en algún intervalo alrededor de x_0 y los q_i no tienen factores comunes. Los puntos dados por cada desarrollo son las llamadas ramas de la curva algebraica.

Las intersecciones de una curva y una recta y de dos curvas es otro tema que recibió gran cantidad de atención. Stirling, en sus *Lineae (Líneas)* de 1717, demostró que una curva algebraica de grado n-ésimo (en x e y) está determinada por $n(n + 3)/2$ de sus puntos, ya que tiene ese número de coeficientes esenciales. También afirmó que dos rectas paralelas cualesquiera cortan a una curva en el mismo número de puntos, reales o imaginarios, y probó que el número de ramas de una curva que se se extiende al infinito es par. La obra de Maclaurin *Geometría orgánica* fundó la teoría de intersecciones de curvas planas superiores: generalizó resultados para casos especiales y sobre esta base concluyó que una ecuación de grado m-ésimo y una de grado n-ésimo se intersecan en mn puntos.

En 1748, Euler y Cramer buscaron demostrar este resultado, pero ninguno ofreció una prueba correcta. Euler [10] se apoyó en un argumento por analogía; dándose cuenta de que su argumento no era completo, dijo que se debía aplicar el método a ejemplos particulares. En su libro de 1750, la «prueba» de Cramer descansaba por completo en ejemplos y, desde luego, no fue aceptada. Ambos tomaron en cuenta los puntos de intersección con coordenadas imaginarias y los que estaban en el infinito, y notaron que el número mn sería obtenido sólo si ambos tipos de puntos eran incluidos y si cualquier factor, tal como $ax + by$, común a ambas curvas era excluido.

---

[9] *Jour. de Math.*, 13, 1850, 365-480.

[10] *Mém. de l'Acad de Berlin*, 4, 1748, 234-248 = *Opera*, (1), 26, 46-59.

Sin embargo, ambos fracasaron al asignar la debida multiplicidad a varios tipos de intersecciones. En 1764, Etienne Bezout (1730-83) dio una prueba mejor del teorema, pero ésta también estaba incompleta en la cuenta de la multiplicidad asignada a puntos en el infinito y puntos múltiples. El cálculo correcto de la multiplicidad fue solucionado por Georges Henri Halphen (1844-89) en 1873 [11].

En su libro de 1750 Cramer estudió una paradoja mencionada por Maclaurin en su *Geometría* concerniente al número de puntos comunes a dos curvas. Una curva de grado n está determinada por $n(n+3)/2$ puntos. Dos curvas de grado n-ésimo se encuentran en n^2 puntos. Ahora, si n es 3, la curva deberá, por el primer argumento, estar determinada por nueve puntos. Así, ya que dos curvas de tercer grado se encuentran en nueve puntos, esos puntos no determinan una curva única de tercer grado. Una paradoja similar surge cuando $n=4$. La explicación de Cramer respecto a esta paradoja, que ahora se le adjudica a él, era que las n^2 ecuaciones que determinan los n^2 puntos de la intersección no son independientes. Todas las cúbicas que pasan por ocho puntos fijos de una cúbica dada deben pasar por el mismo noveno punto fijo. Esto es, el noveno es dependiente de los ocho primeros. Euler proporcionó la misma explicación en 1748 [12].

En 1756 Matthieu B. Goudin (1734-1817) y Achille Pierre Dionis du Séjour (1734-94) escribieron el *Traité des courbes algébriques (Tratado de curvas algebraicas).* Sus nuevas características son que una curva de orden (grado) n no puede tener más de $n(n-1)$ tangentes con una dirección dada, ni más de n asíntotas. Como ya había hecho Maclaurin, ellos señalaron que una asíntota no puede cortar la curva en más de $n-2$ puntos.

En el siglo XVIII, los dos mejores compendios de resultados sobre curvas planas de grado superior son el segundo volumen de la *Introductio* (1748) de Euler y el *Lignes courbes algébriques (Líneas curvas algebraicas)* de Cramer. El último libro posee una gran unidad de punto de vista, está excelentemente expuesto y reseña muy buenos ejemplos. El trabajo era citado con frecuencia, llegando al extremo de adjudicarse a Cramer algunos resultados que originalmente no eran suyos.

---

[11] *Bull. Soc. Math. de France*, 1, 1873, 130-148; 2, 1873, 34-52; 3, 1875, 76-92 = *Œuvres*, 1, 98-157, 171-193 y 337-357.

[12] *Mém. de l'Acad de Berlin*, 4, 1748, 219-233, pub. 1750 = *Opera*, (1), 26, 33-45.

4. Los orígenes de la geometría diferencial

Mientras la geometría diferencial estaba en sus inicios, la geometría analítica se extendía, y de tal forma que los dos desarrollos con frecuencia se entretejieron. El interés por la teoría de las curvas algebraicas declinó durante la última parte del siglo XVIII, y la geometría diferencial se hizo más importante en cuanto concernía a la geometría. Esta materia consiste en el estudio de aquellas propiedades de las curvas y superficies que varían de punto a punto y que, por lo tanto, pueden ser comprendidas únicamente con las técnicas del cálculo. El término «geometría diferencial» se usó por primera vez por Luigi Bianchi (1856-1928) en 1894.

En gran medida, la geometría diferencial fue el crecimiento natural de los propios problemas del cálculo. Las consideraciones sobre las normales a las curvas, puntos de inflexión y curvaturas es de hecho la geometría diferencial de curvas planas. Sin embargo, muchos nuevos problemas de la última parte del siglo XVII y primera del XVIII, el mayor conocimiento acerca de la curvatura de las curvas planas y espaciales, envolventes de familias de curvas, geodésicas sobre superficies, el estudio de los rayos de luz y de superficies de onda de luz, problemas dinámicos de movimiento junto con curvas y restricciones impuestas por superficies, y, por encima de todo, el trazado de mapas llevaron a preguntas sobre curvas y superficies: se hizo evidente que se tenía que aplicar el cálculo.

Los cultivadores de la geometría diferencial del siglo XVIII y primera parte del XIX usaron argumentos geométricos junto con los analíticos, aunque estos últimos dominaron la escena. El análisis era aún basto. El infinitesimal o diferencial de una variable independiente era observado como una constante en extremo pequeña. No se había dado una verdadera distinción entre el incremento de una variable independiente y el diferencial. A las diferenciales de órdenes superiores se las tomaba en cuenta, pero se las veía a todas como muy pequeñas, dándose la libertad de despreciarlas. Los matemáticos hablaban de puntos adyacentes o próximos de una curva como si no hubiera puntos entre los puntos adyacentes si la distancia entre los dos era suficientemente pequeña, por lo que una tangente a una curva conectaba un punto con el siguiente.

5. Curvas planas

La primera aplicación del cálculo a las curvas se hizo con las curvas planas. Algunos de los conceptos subsecuentemente tratados por el cálculo fueron introducidos por Christian Huygens, quien usó únicamente métodos geométricos. Su trabajo en este sentido estuvo motivado por su interés en la luz y el diseño de relojes de péndulo. En 1673, en el tercer capítulo de su *Horologium Oscillatorium (Péndulo oscilatorio),* introdujo la involuta a una curva plana C. Imagínese una cuerda enrollada alrededor de C, de P_1 hacia la derecha (fig. 23.10). El extremo en P_1 de C se mantiene fijo y el otro desenrollado mientras que la cuerda se mantiene tensa. El lugar C' del extremo libre es una evoluta de C. Huygens probó que en el extremo libre la cuerda es perpendicular al foco C'. Cada punto de la cuerda también describe una involuta; por lo que C'' también es una involuta, y Huygens probó que las involutas no se pueden tocar una a la otra. Dado que la cuerda es tangente a C en el punto donde se separa de C, se sigue que cada trayectoria ortogonal de la familia de las tangentes a la curva es una involuta a la curva.

Huygens trató entonces la evoluta de una curva plana. Dada una normal fija en un punto P sobre la curva, cuando una normal adyacente se mueva hacia ella, el punto de intersección de las normales llega a una posición límite sobre la normal fija, el cual es llamado centro de curvatura de la curva en P. Huygens mostró que la distancia desde el punto sobre la curva a lo largo de la normal fija a la posición límite es (en notación moderna)

$$\frac{[1 + (dy/dx]^{3/2}}{d^2y/dx^2}$$

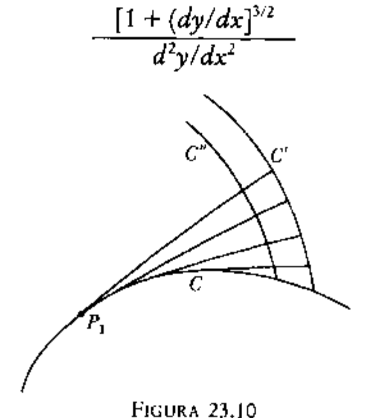


FIGURA 23.10

Esta longitud es el radio de curvatura de la curva en *P*. El foco de los centros de la curvatura, uno para cada normal, es llamado la evoluta de la curva original, por lo que la curva *C* es la evoluta de cualesquiera de sus involutas. En este trabajo demostró Huygens que la evoluta de una cicloide es una cicloide, o más precisamente, la evoluta de la mitad izquierda de la cicloide menor en la figura 23.11 es la mitad derecha de la cicloide superior. Este teorema fue demostrado analíticamente por Euler en 1764 [13]. El significado de la cicloide para el trabajo de Huygens sobre los relojes de péndulo es que un péndulo que se mueve oscilando a lo largo de un arco de cicloide tarda exactamente el mismo tiempo para completar giros de amplitudes cortas y largas. Por esta razón la cicloide es llamada tautócrona.

También Newton, en su *Geometria Analytica* (*Geometría analítica*, publicada en 1736 aunque la mayor parte de ella fue escrita en 1671) introduce los centros de curvatura como los puntos límites de la intersección de una normal en *P* con una normal adyacente. Entonces afirma que el círculo con centro en el centro de curvatura y radio igual al radio de curvatura es el círculo de contacto más cercano a la curva en *P*; es decir, ningún otro círculo tangente a la curva en *P* puede interponerse entre la curva y el círculo de contacto más cercano. Este círculo de contacto más cercano es llamado el círculo osculador, habiendo sido usado el término «osculador» por Leibniz en un ensayo de 1686 [14]. La curvatura de este círculo es el recíproco de su radio y es la curvatura de la curva en *P*. Newton proporcionó

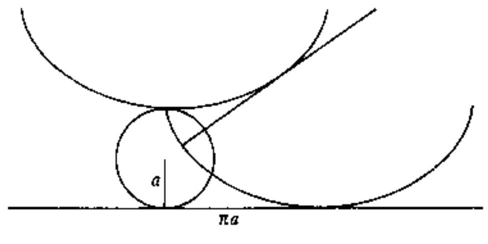


FIGURA 23.11

[13] *Novi Comm. Acad. Sci. Petrop.*, 10, 1764, 179-198, pub. 1766 = *Opera*, (1), 27, 384-400.

[14] *Acta Erud.*, 1686, 289-292 = *Math. Schriften*, 2, 166; 3, 326-329.

también la fórmula para la curvatura y calculó la curvatura de varias curvas, incluyendo la cicloide. Notó que en el punto de inflexión una curva tiene curvatura cero. Estos resultados duplican aquellos de Huygens, pero Newton probablemente deseaba mostrar que él podía utilizar métodos analíticos para establecerlos.

En 1691, Jean Bernoulli retomó el estudio de las curvas planas y obtuvo algunos resultados nuevos sobre envolventes. La cáustica de una familia de rayos de luz, esto es, la envolvente de la familia, había sido introducida por Tschirnhausen en 1682. En el *Acta Eruditorum* de 1692, Bernoulli obtuvo las ecuaciones de algunas cáusticas, por ejemplo, la cáustica de rayos reflejada desde un espejo esférico cuando un haz de rayos paralelos choca con él [15]. Más adelante atacó el problema propuesto por Fatio de Guillier: encontrar la envolvente de la familia de parábolas que son trayectorias de balas de cañón disparadas desde un cañón con la misma velocidad inicial, pero con diferentes ángulos de elevación. Bernoulli mostró que la envolvente es una parábola cuyo foco está en el arma. Este resultado ya había sido establecido geométricamente por Torricelli. En el *Acta Eruditorum* de 1692 y 1694 [16] Leibniz proporcionó el método general para encontrar la envolvente de una familia de curvas. Si la familia está dada (en nuestra notación) por $f(x, y, \alpha) = 0$, donde α es el parámetro de la familia, el método requiere eliminar α entre $f = 0$ y $\partial f/\partial\alpha = 0$. El texto de L'Hospital, *L'analyse des infiniment petits* (*El análisis de los infinitamente pequeños*, 1696) ayudó a perfeccionar y extender la teoría de las curvas planas.

6. Curvas en el espacio

Clairaut inició la teoría de las curvas en el espacio, el primer avance importante en la geometría diferencial de tres dimensiones. Alexis-Claude Clairaut (1713-65) fue precoz. A los doce años de edad ya había escrito un buen trabajo sobre las curvas. En 1731 publicó *Recherche sur les courbes à double courbure (Investigación sobre las curvas de doble curvatura)*, que escribió en 1729, cuando tenía dieciséis años de edad. En este libro discutió el aspecto analítico de las superficies y curvas espaciales (sec. 2). Otro ensayo de

[15] *Opera*, 1, 52-59.

[16] Página 311; véase también *Math. Schriften*, 2, 166; 3, 967 y 969.

Clairaut lo llevó a su elección a la Academia de Ciencias de París a la edad sin precedente de diecisiete años. En 1743 realizó su trabajo clásico sobre la forma de la Tierra. Aquí trató de una manera más completa de lo que lo habían hecho Newton y Maclaurin la forma de un cuerpo que gira, tal como la que toma la Tierra bajo la atracción gravitacional mutua de sus partes. También trabajó sobre el problema de los tres cuerpos, principalmente para estudiar el movimiento de la Luna (cap. 21, sec. 7), escribió varios ensayos sobre ello, y uno por el que ganó el premio de la Academia de San Petersburgo en 1750. En 1763 publicó su *Théorie de la Lune (Teoría de la Luna).* Clairaut poseía un gran encanto personal y fue una figura de la sociedad parisina.

En su trabajo de 1731 trató analíticamente problemas fundamentales de las curvas en el espacio. Llamó a las curvas espaciales «curvas de doble cuadratura» porque, siguiendo a Descartes, consideraba sus proyecciones sobre dos planos perpendiculares: las curvas espaciales entonces «participan» de las curvaturas de las dos curvas sobre los planos. Pensó geométricamente una curva en el espacio como la intersección de dos superficies; analíticamente, la ecuación de cada superficie era expresada como una ecuación de tres variables (sec. 2). Clairaut estudió entonces tangentes a las curvas, y vio que una curva puede tener infinitas normales localizadas en un plano perpendicular a la tangente. Las expresiones para la longitud de arco de una curva espacial y la cuadratura de ciertas áreas sobre superficies también se deben a Clairaut.

A pesar de que Clairaut había dado algunos pasos en la teoría de las curvas en el espacio, muy poco se había hecho en esta materia o en la teoría de las superficies en 1750. Esto se refleja en la *Introductio (Introducción)* de Euler de 1748, donde presentó la geometría diferencial de las figuras planas y espaciales. La primera era bastante completa, pero no así la segunda.

El siguiente paso importante en la geometría diferencial de curvas espaciales fue dado por Euler. Gran parte de su trabajo en geometría diferencial estuvo motivado por uso de curvas y superficies en mecánica. Su *Mechanica* (*Mecánica*, 1736) [17], escrita cuando contaba veintinueve años de edad, es una contribución muy importante a los fundamentos analíticos de la mecánica. Proporcionó otro tratamiento de la materia en su *Theoria Motus Corporum Solidorum seu Ri-*

[17] *Opera*, (2), 1 y 2.

gidorum (*Teoría del movimiento de los cuerpos sólidos y rígidos*, 1765)[18]. En este trabajo obtuvo las fórmulas actuales para el uso de coordenadas polares en las componentes radial y normal de la aceleración de una partícula moviéndose a lo largo de una curva plana, es decir,

$$a_r = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2, \qquad a_\theta = r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}.$$

Empezó a escribir sobre la teoría de las curvas en el espacio en 1774. El problema particular que posiblemente motivó a Euler a abordar esta teoría fue el estudio de la elástica, es decir, la forma tomada por una banda inicialmente recta cuando, bajo presión sobre sus extremos, es doblada y torcida en la forma de una curva alabeada. Para tratar este problema introdujo algunos nuevos conceptos en 1774[19]. Más adelante proporcionó un tratamiento completo de la teoría de estas curvas en un ensayo presentado en 1775[20].

Euler representó las curvas espaciales mediante las ecuaciones paramétricas $x = x(s)$, $y = y(s)$, $z = z(s)$, donde s es la longitud de arco, y, como otros autores del siglo XVIII, utilizó la trigonometría esférica para llevar a cabo su análisis. A partir de las ecuaciones paramétricas obtiene

$$dx = p\ ds, \qquad dy = q\ ds, \qquad dz = r\ ds,$$

donde p, q y r son cosenos directores, variando de punto a punto y, por supuesto, con $p^2 + q^2 + r^2 = 1$. La cantidad ds, diferencial de la variable independiente, la consideraba como una constante.

Para estudiar las propiedades de la curva introdujo la indicatriz esférica. Alrededor del punto (x, y, z) de la curva, Euler describe una esfera de radio 1. La indicatriz esférica puede ser definida como el lugar sobre la esfera unitaria de los puntos cuyos vectores de posición partiendo del centro O son iguales a la tangente unitaria en (x, y, z) y las tangentes unitarias en los puntos vecinos. De ahí que los dos radios en la figura 23.12 representen una tangente unitaria

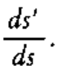
---

18 *Opera*, (2), 3 y 4.

19 *Novi Comm. Acad. Sci. Petrop.*, 19, 1774, 340-370, pub. 1775 = *Opera*, (2), 11, 158-179.

20 *Acta Acad. Sci. Petrop.*, 1, 1782, 19-57, pub. 1786 = *Opera*, (1), 28, 348-381.

en (x, y, z) y el punto vecino de la curva. Sea ds' el arco o el ángulo entre dos tangentes vecinas de los dos puntos que están aquí separados ds a lo largo de la curva. La definición de Euler del radio de curvatura de la curva es



$$\frac{ds'}{ds}.$$



FIGURA 23.12

Más adelante deriva una expresión analítica del radio de curvatura

$$\varrho = \frac{ds^2}{\sqrt{(d^2 x)^2 + (d^2 y)^2 + (d^2 z)^2}} \left[= \frac{1}{\sqrt{x''^2 + y''^2 + z''^2}} \right]. \quad (8)$$

El plano que contiene ds' y el centro O es la definición de Euler del plano osculador en (x, y, z). Jean Bernoulli, quien introdujo el término, consideraba el plano como determinado por tres puntos «coincidentes». Su ecuación, tal como fue dada por Euler, es

$$x(r\,dq - q\,dr) + y(p\,dr - r\,dp) + z(q\,dp - p\,dq) = t,$$

donde t está determinada por el punto (x, y, z) sobre la curva por el que pasa el plano osculador. Esta ecuación es equivalente a la que se escribe en notación vectorial hoy en día como

$$(\mathbf{R} - \mathbf{r}) \cdot \mathbf{r}' \times \mathbf{r}'' = 0,$$

donde $\mathbf{r}(s)$ es el vector de posición con respecto a algún punto en el espacio del punto sobre la curva en el cual el plano osculador está determinado y $\mathbf{R}$ es el vector de posición en cualquier punto en el plano oscilante. En forma vectorial, $\mathbf{r}$ está dado por

$$x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k}$$

y **R** tiene la forma $X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$, donde (Y, Y, Z) son las coordenadas de **R**.

Clairaut había introducido la idea de que una curva espacial tiene dos curvaturas. Una de éstas fue sistematizada por Euler en la forma que acabamos de describir. La otra, ahora llamada «torsión» y que representa geométricamente la proporción en la que una curva se aleja de un plano en un punto (x, y, z), fue formulada explícita y analíticamente por Michel Ange Lancret (1774-1807), ingeniero y matemático alumno de Monge, quien trabajó en esa tradición. Lancret señaló [21] tres direcciones principales en cualquier punto sobre una curva. El primero es el de la tangente. Las tangentes «sucesivas» descansan en el plano osculador. La normal a la curva que descansa en el plano osculador es la normal principal, y la perpendicular al plano osculador, la binormal, es la tercera dirección principal. La torsión es la proporción de cambio en la dirección de la binormal con respecto a la longitud de arco. Lancret usó la terminología flexión de planos osculadores sucesivos o binormales sucesivas.

Lancret representó una curva mediante

$$x = \phi(z), \qquad y = \psi(z)$$

y llamó $d\mu$ el ángulo entre los planos normales sucesivos y $d\omega$ el correspondiente entre los osculadores sucesivos. Entonces, en notación moderna

$$\frac{d\mu}{ds} = \frac{1}{\varrho}, \qquad \frac{d\nu}{ds} = \frac{1}{\tau},$$

donde ϱ es el radio de curvatura y τ es el radio de torsión.

Cauchy mejoró la formulación de los conceptos y precisó gran parte de la teoría de curvas espaciales en su famoso *Leçons sur les applications du calcul infinitesimal à la géometrie (Lecciones sobre las aplicaciones del cálculo infinitesimal a la geometría*, 1826) [22]. Descartó los infinitesimales constantes, los ds', y aclaró la confusión entre los incrementos y las diferenciales. Señaló que cuando alguien escribe

$$ds^2 = dx^2 + dy^2 + dz^2$$

---

[21] *Mém. divers Savans*, 1, 1806, 416-454.
[22] *Œuvres*, (2), 5.

se debe entender

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2.$$

Cauchy prefería escribir una superficie como $w(x, y, z) = 0$ en vez de la forma asimétrica $z = f(x, y)$, y escribía la ecuación de una línea recta pasando por el punto (ξ, η, ζ) como

$$\frac{\xi - x}{\cos\alpha} = \frac{\eta - y}{\cos\beta} = \frac{\zeta - z}{\cos\gamma},$$

donde $\cos\alpha$, $\cos\beta$ y $\cos\gamma$ son los cosenos directores de la recta, aunque utilizó más a menudo números directores en lugar de cosenos directores.

Prácticamente, el desarrollo de Cauchy de la geometría de curvas es moderno. Eliminó la trigonometría esférica de sus demostraciones, pero también tomó la longitud de arco como variable independiente y obtiene para los cosenos directores de la tangente en cualquier punto

$$\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}, \quad \text{o} \quad x'(s), y'(s), z'(s).$$

Los números directores de la normal principal se muestra que son

$$\frac{d^2x}{ds^2}, \frac{d^2y}{ds^2}, \frac{d^2z}{ds^2}, \quad \text{o} \quad x''(s), y''(s), z''(s)$$

y la curvatura k de la curva es

$$k = \frac{1}{\varrho} = \sqrt{(x'')^2 + (y'')^2 + (z'')^2}.$$

Entonces, demuestra que si los cosenos directores de la tangente son $\cos\alpha$, $\cos\beta$ y $\cos\gamma$

$$x'' = \frac{d(\cos\alpha)}{ds} = \frac{\cos\lambda}{\varrho}, \qquad y'' = \frac{d(\cos\beta)}{ds} = \frac{\cos\mu}{\varrho},$$

$$z'' = \frac{d(\cos\nu)}{ds} = \frac{\cos\gamma}{\varrho}, \tag{9}$$

donde ϱ es el radio de curvatura ya introducido y cos λ, cos μ y cos ν son los cosenos directores de una normal, que él toma como la principal. Enseguida demuestra que

$$\frac{1}{\varrho} = \frac{d\omega}{ds},$$

donde ω es el ángulo entre las tangentes adyacentes.

Introduce el plano osculador como el plano de la tangente y la normal principal. La normal a este plano es la binormal y sus cosenos directores cos L, cos M y cos N están dados por las fórmulas

$$\frac{\cos L}{dy\, d^2z - dz\, d^2y} = \frac{\cos M}{dz\, d^2x - dx\, d^2z} = \frac{\cos N}{dx\, d^2y - dy\, d^2x}.$$

Entonces puede probar que

$$\frac{d \cos L}{ds} = \frac{\cos \lambda}{\tau}, \qquad \frac{d \cos M}{ds} = \frac{\cos \mu}{\tau},$$

$$\frac{d \cos N}{ds} = \frac{\cos \nu}{\tau}, \tag{10}$$

donde $1/\tau$ es la torsión, la cual es igual a $d\Omega/ds$, donde Ω es el ángulo entre los planos osculadores.

Las fórmulas (9) y (10) son dos de las tres famosas fórmulas de Serret-Frénet, la tercera es

$$\frac{d \cos \lambda}{ds} = -\frac{\cos \alpha}{\varrho} - \frac{\cos L}{\tau},$$

$$\frac{d \cos \mu}{ds} = -\frac{\cos \beta}{\varrho} - \frac{\cos M}{\tau},$$

$$\frac{d \cos \nu}{ds} = -\frac{\cos \gamma}{\varrho} - \frac{\cos N}{\tau}, \tag{11}$$

donde $1/\tau$ es la torsión y $1/\varrho$ es la curvatura. Estas fórmulas (9), (10) y (11), que proporcionan las derivadas de los cosenos directores de la tangente, binormal y normal, respectivamente, fueron publicadas por Joseph Alfred Serret (1819-85) en 1851 [23] y Fréderic-Jean Frénet

[23] *Jour. de Math.*, 16, 1851, 193-207.

(1816-1900) en 1852 [24]. La significación de la curvatura y la torsión es que son dos propiedades esenciales de las curvas en el espacio. Dadas la curvatura y torsión como funciones de la longitud del arco a lo largo de la curva, ésta está completamente determinada excepto por su posición en el espacio. Este teorema es fácilmente demostrable a partir de las fórmulas de Serret-Frénet.

7. La teoría de superficies

Como la teoría de curvas en el espacio, la teoría de superficies tuvo un inicio lento. Empezó con el estudio de las geodésicas sobre superficies, con las geodésicas sobre la Tierra como preocupación principal. En el *Journal des Scavans* de 1697, Jean Bernoulli propuso el problema de encontrar el arco mínimo entre dos puntos sobre una superficie convexa [25]. Le escribió a Leibniz en 1698 para señalarle que el plano osculador (el plano del círculo osculador) en cualquier punto de una geodésica es perpendicular a la superficie en ese punto. En 1698 Jacques Bernoulli resolvió el problema de las geodésicas sobre cilindros, conos y superficies de revolución. El método era limitado, a pesar de que en 1728 Jean Bernoulli [26] tuvo cierto éxito con el método y encontró las geodésicas sobre otros tipos de superficies.

En 1728 Euler [27] proporcionó las ecuaciones diferenciales para las geodésicas sobre superficies. Euler usó el método que había introducido en el cálculo de variaciones (véase cap. 24, sec. 2). En 1732 Jacob Hermann [28] también encontró geodésicas en superficies particulares.

En 1733 y también en 1739 [29], en su trabajo sobre la forma de la Tierra, Clairaut trató las geodésicas sobre superficies de revolución de una manera más completa. Demostró en su ensayo de 1733 que para cualquier superficie de revolución, el seno del ángulo formado por una curva geodésica y cualquier meridiano (cualquier po-

[24] *Jour. de Math.*, 17, 1852, 437-447.

[25] *Opera*, 1, 204-205.

[26] *Opera*, 4, 108-128.

[27] *Comm. Acad. Sci. Petrop.*, 3, 1728, 110-124, pub. 1732 = *Opera*, (1), 25, 1-12.

[28] *Comm. Acad. Sci. Petrop.*, 6, 1732/3, 36-67.

[29] *Hist. de l'Acad. des Sci., Paris*, 1733, 186-194, pub. 1735 y 1739, 83-96, pub. 1741.

sición de la curva generadora) que cruza, varía inversamente con la longitud de la perpendicular a partir del punto de intersección al eje. En otro ensayo [30], demostró también el bello teorema de que si en cualquier punto M de una superficie de revolución un plano pasa por la normal a la superficie y al plano del meridiano en M, entonces la curva determinada en la superficie tiene un radio de curvatura en M igual a la longitud de la normal entre M y el eje de revolución. Los métodos de Clairaut eran analíticos, pero, como la mayoría de sus predecesores, no utilizó las ideas que hoy en día asociamos con el cálculo de variaciones.

En 1760, en su *Recherches sur la courbure des surfaces (Investigaciones sobre la curvatura de las superficies)* [31], Euler estableció la teoría de las superficies. Este trabajo es la contribución más importante de Euler a la geometría diferencial y un punto culminante de la materia. Euler representa una superficie por $z = f(x, y)$ e introduce la actual notación

$$p = \frac{\partial z}{\partial x}, \qquad q = \frac{\partial z}{\partial y}, \qquad r = \frac{\partial^2 z}{\partial x^2}, \qquad s = \frac{\partial^2 z}{\partial x\, \partial y}, \qquad t = \frac{\partial^2 z}{\partial x^2}.$$

Más adelante dice: «Empiezo por determinar el radio de curvatura de cualquier sección plana de una superficie; entonces aplico esta solución a secciones que son perpendiculares a la superficie en cualquier punto dado; finalmente, comparo los radios de curvatura de estas secciones con respecto a su mutua inclinación, lo que nos coloca en situación de establecer una idea adecuada de la curvatura de superficies.»

Primero obtiene una expresión bastante compleja del radio de curvatura de cualquier curva obtenida al cortar una superficie con un plano. Entonces particulariza el resultado al aplicarlo a secciones normales (secciones conteniendo una normal a una superficie). Para las secciones normales, la expresión general para el radio de curvatura se simplifica un poco. Más adelante define la sección normal principal como la sección normal perpendicular al plano xy. (Este uso de «principal» no se sigue hoy en día.) El radio de curvatura de una sección normal cuyo plano forma un ángulo ϕ con el plano de

---

30 *Mém. de l'Acad. des Sci., Paris*, 1735, 117-122, pub. 1738.

31 *Mém. de l'Acad. de Berlin*, 16, 1760, 119-143, pub. 1767 = *Opera*, (1), 28, 1-22.

la sección normal principal tiene la forma

$$\frac{1}{L + M\cos 2\phi + N\,\mathrm{sen}\,2\phi}\,, \tag{12}$$

donde L, M y N son funciones de x e y. Para obtener la máxima y mínima curvaturas de todas las secciones normales en un punto sobre la superficie (o cuando la forma del denominador en (12) es indefinida, para obtener las dos curvaturas mayores), hace la derivada con respecto a ϕ del denominador igual a cero y (en ambos casos) obtiene tg $2\phi = N/M$. Existen dos raíces que difieren en 90°, de tal manera que hay dos planos normales mutuamente perpendiculares. Llamamos a las curvaturas correspondientes las curvaturas principales $\varkappa_1$ y $\varkappa_2$.

Se sigue de los resultados de Euler que la curvatura de cualquier otra sección normal que forma un ángulo α con una de las secciones con curvatura principal es

$$\varkappa = \varkappa_1 \cos^2 \alpha + \varkappa_2\,\mathrm{sen}^2\,\alpha. \tag{13}$$

Este resultado se llama teorema de Euler.

Los mismos resultados fueron obtenidos en 1776 de una manera más elegante por un alumno de Monge, Jean-Baptiste-Marie-Charles Meusnier de la Place (1754-93), quien también trabajó en hidrostática y en química con Lavoisier. Más adelante, Meusnier [32] estudió la curvatura de secciones no normales, para las cuales Euler había obtenido una expresión muy compleja. El resultado de Meusnier, aún llamado teorema de Meusnier, afirma que la curvatura de una sección plana de una superficie en un punto P es la curvatura de la sección normal a lo largo de la misma tangente en P dividida por el seno del ángulo que el plano original forma con el plano tangente en P. Le sigue el bello resultado de que si uno considera la familia de planos a lo largo de la misma línea tangente MM' a una superficie, entonces los centros de curvatura de las secciones determinadas por estos planos están en un círculo en un plano perpendicular a MM' y cuyo diámetro es el radio de curvatura de la sección normal. Meusnier, entonces, prueba el teorema de que las únicas superficies para las cuales las dos curvaturas principales son iguales en todo punto son los planos o las esferas. Su ensayo fue particularmente

[32] *Mém. divers Savans*, 10, 1785, 477-485.

simple y fértil; ayudó a hacer intuitivos un buen número de resultados obtenidos durante el siglo XVIII.

Una preocupación especial de la teoría de las superficies, motivada por las necesidades de la cartografía, es el estudio de las superficies desarrollables: dado que la esfera no puede ser cortada y luego aplanada, el problema era encontrar superficies con formas cercanas a la de una esfera que pudieran ser desarrolladas sin gran distorsión. Euler fue el primero en considerar el tema. Este trabajo se encuentra contenido en su *«De Solidis Quorum Superficiem in Planum Explicare Licet»* [33] *(Los sólidos...)*. En el siglo XVIII, las superficies eran consideradas como fronteras de sólidos; por ello habla Euler de sólidos cuyas superficies pueden ser desdobladas sobre un plano. En este ensayo introduce la representación paramétrica de superficies, esto es:

$$x = x(t, u), \qquad y = y(t, u), \qquad z = z(t, u),$$

y pregunta qué condiciones deben satisfacer estas funciones de tal manera que la superficie pueda ser desarrollable sobre un plano. Su método es representar t y u como coordenadas rectangulares en un plano y entonces formar un pequeño triángulo rectángulo (t, u), $(t + dt, u)$, $(t, u + du)$. Ya que la superficie es desarrollable, este triángulo debe ser congruente con un triángulo pequeño sobre la superficie. Si denotamos las derivadas parciales de x, y y z con respecto a t por l, m y n, y las derivadas con respecto a u por λ, μ y ν, entonces el triángulo correspondiente en la superficie es (x, y, z), $(x + l\,dt, y + m\,dt, z + n\,dt)$ y $(x + \lambda\,du, y + \mu\,du, z + \nu\,du)$. Como consecuencia de la congruencia entre los dos triángulos, Euler obtiene

$$l^2 + m^2 + n^2 = 1, \qquad \lambda^2 + \mu^2 + \nu^2 = 1, \qquad l\lambda + m\mu + n\nu = 0. \qquad (14)$$

Estas son las condiciones analíticas necesarias y suficientes para la desarrollabilidad. La condición es equivalente a requerir que el elemento de curva sobre la superficie sea igual al elemento de línea sobre el plano. Analíticamente, el problema de determinar si una superficie es desarrollable sería encontrar una representación paramétrica $x(t, u)$, $y(t, u)$ y $z(t, u)$ tal que las derivadas parciales satisfagan las condiciones (14).

---

[33] *Novi Comm. Acad. Sci. Petrop.*, 16, 1771, 3-34, pub. 1772 = *Opera*, (1), 28, 161-186.

Más adelante, Euler investigó la relación entre las curvas en el espacio y las superficies desarrollables y mostró que la familia de tangentes a cualquier curva en el espacio llena o constituye una superficie desarrollable. Intentó sin éxito demostrar que toda superficie desarrollable es una superficie reglada, esto es, una superficie generada por una línea recta en movimiento, y a la inversa. El recíproco es, de hecho, falso.

Gaspard Monge trató independientemente el tema de las superficies desarrollables. Con Monge, la geometría y el análisis se apoyaban mutuamente. Consideró simultáneamente ambos aspectos del mismo problema y mostró la utilidad de pensar tanto geométrica como analíticamente. El efecto del doble punto de vista de Monge fue colocar a la geometría, al menos, en igualdad con el análisis e inspirar el renacer de la geometría pura. A pesar de que el análisis había dominado en el siglo XVIII, a pesar de la geometría analítica y la geometría diferencial de Euler y Clairaut, Monge es el primer y verdadero innovador en geometría sintética después de Desargues.

Su extenso trabajo en geometría descriptiva (la cual sirve principalmente a la arquitectura), geometría analítica, geometría diferencial y ecuaciones diferenciales ordinarias y parciales ganó la admiración y envidia de Lagrange, quien después de escuchar una conferencia de Monge, le dijo: «Usted acaba de presentar, mi querido colega, muchas cosas elegantes. Yo desearía haberlas hecho.» Monge contribuyó mucho a la física, química, metalurgia (problemas de forja) y maquinaria. En química trabajó con Lavoisier. Monge vio la necesidad de la ciencia en el desarrollo de la industria y abogó por la industrialización como una vía por la mejora de la vida. Estaba inspirado por una activa preocupación social, tal vez porque conoció el infortunio de orígenes humildes; por esta razón apoyó la Revolución Francesa y sirvió como ministro de la Marina y como miembro del Comité de Salvación Pública en los gobiernos sucesivos. Diseñó armamentos e instruyó a personal del gobierno en asuntos técnicos. Su admiración por Bonaparte lo sedujo a seguirlo con sus medidas antirrevolucionarias.

Monge ayudó a organizar la Ecole Polytechnique y como profesor fundó una escuela de geómetras. Fue un gran maestro y una fuerza inspiradora de la actividad matemática del siglo XIX. Sus conferencias vigorosas y fértiles comunicaron entusiasmo a sus alumnos, entre quienes estuvieron al menos una docena de los más famosos personajes del siglo XIX.

A él se deben resultados en geometría diferencial tridimensional que van mucho más allá que los de Euler. Su artículo de 1771 «Mémoire sur les développées, les rayons de courbure et les différents genres d'inflexions des courbes à double courbure» («Memoria sobre las desarrolladas, los radios de curvatura y los diferentes géneros de inflexión de las curvas de doble curvatura»), publicado mucho después [34], fue seguido por su *Feuilles d'analyse appliquée à la géométrie* (*Hojas de análisis aplicado a la geometría*, 1795; segunda edición, 1801). Las *Feuilles* era tanto geometría diferencial como geometría analítica y ecuaciones diferenciales parciales. Basándose en notas de conferencias ofrece una sistematización y extensión de antiguos resultados, nuevos resultados de cierta importancia y el traslado de varias propiedades de curvas y superficies al lenguaje de las ecuaciones diferenciales parciales. Al perseguir la relación entre las ideas del análisis y las de la geometría, Monge reconoció que una familia de superficies, al tener una propiedad geométrica común o al estar definidas por el mismo método de generación, deberían satisfacer una ecuación diferencial parcial (cap. 22, sec. 6).

El primer trabajo importante de Monge, el ensayo en torno a las superficies desarrollables sobre curvas de doble curvatura, relaciona curvas en el espacio y superficies asociadas a ellas. En ese momento, Monge no conocía el trabajo de Euler sobre las superficies desarrollables. Trata las curvas en el espacio ya sea como la intersección de dos superficies o por medio de sus proyecciones sobre dos planos perpendiculares, dados por $y = \phi(x)$ y $z = \psi(x)$. En cualquier punto ordinario hay una infinidad de normales (perpendiculares a la línea tangente) que están en un plano, el plano normal. En este plano existe una línea que llama eje (polar), el cual es el límite de la intersección con un plano normal vecino. Cuando uno se mueve a lo largo de la curva los planos normales envuelven una superficie desarrollable, llamada la polar desarrollable. La superficie también es engendrada por los ejes de los planos normales. La perpendicular desde P al eje en el plano normal en P es la normal principal y el pie Q de la perpendicular es el centro de curvatura.

Para obtener la ecuación de la polar desarrollable, Euler encuentra la ecuación del plano normal y elimina x entre esta ecuación y la derivada parcial de la ecuación con respecto a x. También obtiene una regla para encontrar la envolvente de cualquier familia de planos

---

[34] *Mém. divers Savans*, 10, 1785, 511-550.

uniparamétrica, que es la que actualmente usamos. La regla, que se aplica igual a familias uniparamétricas de superficies, es esta: en nuestra notación, toma $F(x, y, z, \alpha) = 0$ como la familia. Para encontrar dónde ésta encuentra «una superficie infinitesimalmente cerca», halla $\partial F/\partial\alpha = 0$. La curva de intersección de las dos superficies es llamada la curva característica. La ecuación de la envolvente se encuentra al eliminar α entre $F = 0$ y $\partial F/\partial\alpha = 0$. Euler aplica este método al estudio de otras superficies desarrollables, cada una de las cuales considera como la envolvente de una familia uniparamétrica de planos.

Monge también tuvo en cuenta la arista de regresión *(arête de rebroussement)* de una superficie desarrollable. La arista (curva) está formada por un conjunto de líneas generando la superficie. La intersección de dos líneas vecinas cualquiera es un punto y el lugar de tales puntos es la arista de regresión. Entonces las tangentes a Γ son las generatrices, o Γ es la envolvente de la familia de líneas generatrices. La arista de regresión separa la superficie desarrollable en dos películas u hojas, de la misma manera que una cúspide separa a una curva plana en dos partes. Monge obtuvo las ecuaciones de la arista de regresión. En el caso de la superficie polar desarrollable de una curva en el espacio, la arista de regresión es el lugar geométrico de los centros de curvatura de la curva original.

En 1775, Monge presentó a la Académie des Sciences otro ensayo sobre superficies, particularmente las superficies desarrollables que aparecen en la teoría de las sombras y las penumbras [35]. Usando la definición de que una superficie desarrollable puede ser aplanada sin distorsión, argumenta intuitivamente que una superficie desarrollable es una superficie reglada (pero no recíprocamente) sobre la cual dos líneas consecutivas son concurrentes o paralelas y que cualquier desarrollable es equivalente a aquella formada por las tangentes de una curva en el espacio. Es en este ensayo de 1775 donde proporciona la representación general de las superficies desarrollables. Sus ecuaciones son siempre de la forma

$$z = x[F(q) - qF'(q)] + f(q) - qf'(q),$$

donde $q = \partial z/\partial y$ y la ecuación diferencial parcial que tales superficies, excepto cilindros perpendiculares al plano xy, satisfacen siempre es

$$z_{xx}z_{yy} - z_{xy}^2 = 0.$$

[35] *Mém. divers Savans*, 9, 1780, 382-440.

Después estudia Monge las superficies regladas y da una representación general para ellas. También proporciona la ecuación diferencial parcial de tercer orden que satisfacen, e integra esta ecuación diferencial. Demuestra que las superficies desarrollables son un caso especial de las superficies regladas.

En su «Mémorie sur la théorie des deblais et des remblais» («Memoria sobre la teoría de excavaciones y rellenos») de 1776 [36], considera el problema surgido en la construcción de fortificaciones, que involucra el transporte de tierra y otros materiales de un lugar a otro, y busca encontrar la solución más eficiente, esto es, que el material transportado por la distancia transportada debe ser un mínimo. Unicamente parte de este trabajo es importante para la geometría diferencial de superficies; ciertamente los resultados del problema práctico no son muy realistas —según dice Monge— y publica el ensayo por los resultados geométricos encontrados allí. Aquí inicia el tratamiento del problema de una familia de líneas dependiendo de dos parámetros o de una congruencia de rectas. Siguiendo el trabajo de Euler y Meusnier, considera la familia de normales a una superficie *S*, que forman también una congruencia de líneas. Considérense, en particular, las normales a lo largo de una línea de curvatura (una línea de curvatura es una curva sobre la superficie que tiene una curvatura principal en cada punto de la curva). Las normales a la superficie en estos puntos de la línea de curvatura forman una superficie desarrollable llamada normal desarrollable. Análogamente las normales a la superficie a lo largo de la línea de curvatura perpendicular a la primera forman una superficie desarrollable. Ya que hay dos familias de líneas de curvatura en la superficie, existen dos familias de superficies desarrollables y los miembros de una intersectan a los miembros de la otra en ángulo recto. De hecho, la intersección de dos cualesquiera tiene lugar sobre la normal a la superficie. En cada normal hay dos puntos que distan de la superficie precisamente las dos curvaturas principales. Cada conjunto de puntos determinados por una de las curvaturas principales está en la arista de regresión de una superficie desarrollable formada por ese conjunto de normales, de tal modo que las normales son tangentes a la arista de regresión. Las aristas de regresión de una familia de superficies desarrollables constituyen una superficie llamada superficie central. Por tanto, hay dos superficies centrales (fig. 23.13) como

[36] *Mém. de l'Acad. des Sci., Paris*, 1781, 666-704, pub. 1784.

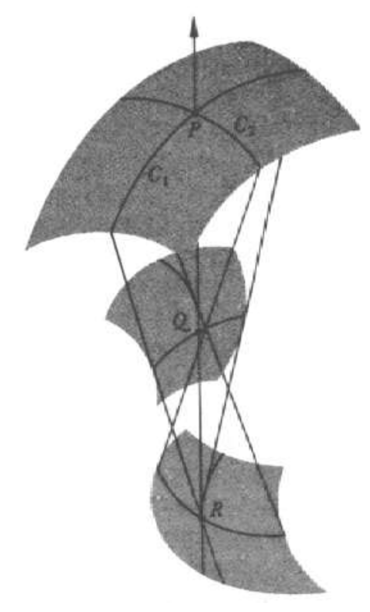


FIGURA 23.13

las descritas. La envolvente de cada familia de superficies desarrollables es llamda superficie focal.

Para el estudio de las ecuaciones diferenciales parciales son muy significativos los detalles ulteriores del trabajo de Monge sobre familias de superficies que satisfacen dichas ecuaciones diferenciales parciales no lineales y lineales de primero, segundo y hasta tercer orden.

Monge explicó sus ideas tratando a fondo un cierto número de curvas y superficies concretas. La generalización y explotación de sus ideas se llevó a cabo por los matemáticos del siglo XIX. Siempre orientado desde un punto de vista práctico, Monge concluyó sus *Feuilles* con una visión de cómo su teoría podía ser aplicada a la arquitectura y, en particular, a la construcción de amplios salones de juntas.

Charles Dupin (1784-1873), alumno de Monge, realizó algunas contribuciones adicionales a la teoría de superficies. Dupin se graduó de la Ecole Polytechnique como ingeniero naval, y, al igual que Monge, tuvo constantemente las aplicaciones geométricas ante él. Su texto *Développements de géométrie* (*Desarrollos de la geometría*,

1813) fue subtitulado «Con aplicaciones a la estabilidad de barcos, excavación y relleno, fortificaciones, óptica, etcétera», y otros escritos, notablemente sus *Applications de géométrie et de mécanique* (*Aplicaciones de la geometría y la mecánica*, 1822), contienen contribuciones a estas materias. Los primeros resultados que describiremos corresponden a su libro de 1813.

Una de las contribuciones de Dupin, que extiende y clarifica trabajos previos de Euler y Meusnier, es la llamada indicatriz de Dupin. Dado el plano tangente a una superficie en un punto M, llevó en cada dirección a partir de M un segmento cuya longitud es igual a la raíz cuadrada del radio de la curvatura de la sección normal de la superficie en esa dirección. El lugar geométrico de los puntos finales de estos segmentos es una cónica, la indicatriz, la cual da una primera aproximación de la forma de la superficie alrededor de M (esto es casi igual a una sección de la superficie hecha por un plano cerca de M y paralelo al plano tangente en M). Las líneas de curvatura de una superficie en el punto, esto es, las curvas sobre la superficie que pasan por M teniendo curvatura extrema (máxima o mínima) son las curvas que tienen como tangentes en M los ejes de la indicatriz.

En un sistema tridimensional de coordenadas rectangulares, las superficies coordenadas son las tres familias de superficies x = const., y = const. y z = const. Supóngase que uno tiene tres familias de superficies, cada una dada por ecuaciones en x, y y z. Si cada miembro de una familia corta ortogonalmente a uno de los miembros de las otras dos familias, entonces las tres familias son llamadas ortogonales. El resultado más sobresaliente de Dupin en esta materia, que se encuentra en su *Développements*, es el teorema de que tres familias de superficies ortogonales se cortan entre sí a lo largo de las líneas de curvatura (las curvas de curvatura normal máxima o mínima) de cada superficie.

Dupin también logró extensiones de los resultados de Monge a las congruencias de líneas. Si la congruencia, la familia biparamétrica, es cortada ortogonalmente por una familia de superficies, como en el caso de la óptica donde la líneas son rayos de luz y las superficies frentes de onda, entonces la congruencia se llama normal. Usando aparentemente los resultados de Monge —aunque no se refiere a ellos— Etienne Louis Malus (1755-1812), físico francés, probó [37]

[37] *Jour. de l'Ecole Poly.*, cahier 14, 1808, 1-44, 84-129.

que una congruencia normal de líneas emanando de un punto (un conjunto homocéntrico) permanece igual después de la reflexión o refracción (de acuerdo con las leyes de la óptica) sobre una superficie. En 1816 [38], Dupin demostró que dicho resultado es cierto para cualquier congruencia normal después de cualquier número de reflecciones. Lambert A. J. Quetelet (1796-1874) proporcionó más adelante una prueba de que una congruencia normal permanece normal después de cualquier número de refracciones [39]. El tema de las congruencias de líneas y complejos de líneas, familias introducidas por Malus y que dependen de tres parámetros, fue seguido por muchos cultivadores en el siglo XIX.

8. El problema de los mapas

Gran parte de la geometría diferencial del siglo XVIII estuvo motivada por problemas de geodesia y trazado de mapas. Sin embargo, el problema de hacer mapas supone consideraciones y dificultades particulares que generaron varios desarrollos matemáticos, en especial transformaciones conformes o aplicaciones conformes, afectando a varias ramas de las matemáticas. El trazado de mapas es, por supuesto, mucho más antiguo que la geometría diferencial, y cada uno de los métodos matemáticos de representación va mucho más atrás. De éstos, la proyección estereográfica y otros métodos (cap. 7, sec. 5) surgen de Ptolomeo y la proyección de Mercator (cap. 12, sec. 2) data del siglo XVI. Respecto a que no eran reales los mapas de una esfera desdoblada en un plano, se pudo haber estado intuitivamente convencido con anterioridad al siglo XVIII, ya que no es posible aplicar una esfera sobre un plano y conservar a la vez las longitudes. Si esto fuera factible, entonces todas las propiedades geométricas podrían ser conservadas. Unicamente las superficies desarrollables pueden ser aplicadas así y, como fue revelado por la aportación del siglo XVIII, éstas son cilindros (no necesariamente circulares), conos y cualquier superficie generada por las líneas tangentes a una curva en el espacio. Puesto que en un mapa de una esfera sobre un plano no es posible conservar todas sus propiedades geométricas,

[38] *Annales de Chimie et de Physique*, 5, 1817, 85-88. También en su *Applications* (*Aplicaciones*) de 1882, 195-197.

[39]. *Correspondance mathématique et physique*, 1, 1825, 147-149.

la atención se dirigió hacia los mapas que conservaban los ángulos.

En tales mapas, si dos curvas sobre una superficie se cortan con un ángulo α y las curvas correspondientes en el otro se cortan con el mismo ángulo, y si la dirección de los ángulos se conserva, se dice que el mapa es conforme. La proyección estereográfica y la proyección de Mercator son conformes. Esta propiedad no significa que dos figuras *finitas* correspondientes sean semejantes, porque la igualdad de ángulos es una propiedad que se tiene en un punto.

J. H. Lambert inició una nueva época en la cartografía teórica, siendo el primero en considerar la aplicación conforme de una esfera sobre el plano en toda su generalidad; en un libro de 1772, *Anmerkungen und Zusatze zur Entwefung der Land-und Himmelscharten (Notas y adiciones sobre diseño de mapas terrestres y mapas de los cielos)*, obtuvo sus fórmulas para esta representación. También Euler hizo muchas contribuciones en este tema y de hecho realizó un mapa de Rusia. En un ensayo presentado a la Academia de San Petersburgo en 1768 [40], Euler, empleando funciones complejas, diseñó un método para representar transformaciones conformes de un plano en otro, pero no lo aprovechó. Más adelante, en dos ensayos presentados en 1775 [41], mostró que la esfera no podía ser aplicada congruentemente en un plano. Aquí también usó funciones complejas, dando un estudio bastante general de las representaciones conformes. Asimismo, proporcionó un análisis completo de las proyecciones de Mercator y de las estereográficas. Lagrange, en 1779 [42], obtuvo todas las transformaciones conformes de una porción de la superficie de la Tierra sobre un área plana, que transformaba círculos de latitud y longitud en arcos circulares.

La difusión de la geometría diferencial y la teoría de funciones complejas tuvo que esperar a que se realizara un mayor progreso en el problema de los mapas y en la representación conforme.

---

[40] *Novi Comm. Acad. Sci. Petrop.*, 14, 1769, 104-128, pub. 1770 = *Opera*, (1), 28, 99-119.

[41] *Acta Acad. Sci. Petrop.*, 1, 1777, 107-132 y 133-142, pub. 1778 = *Opera*, (1), 28, 248-275 y 276-287.

[42] *Nouv. Mém. de l'Acad de Berlin*, 1779, 161-210, pub. 1781 = *Œuvres*, 4, 637-692.

Bibliografía

Ball, W. W. Rouse: «On Newton's classification of Cubic Curves», *Proceedings of the London Mathematical Society*, 22, 1890, 104-143.

Bernoulli, Jean: *Opera Omnia*, 4 vols. 1742, reimpresa por Georg Olms, 1968.

Berzolari, Luigi: «Allgemeine Theorie der hoheren ebenen algebraischen Kurven», *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-15, III C4, 313-455.

Boyer, Carl B.: *History of Analytic Geometry*, Scripta Mathematica, 1956, caps. 6-8.

—: *Historia de la matemática*. Madrid, Alianza, 1986.

Brill, A. y Max Noether: «Die Entwicklung der Theorie der algebraischen Funktionen in älterer und neuerer Zeit», *Jahres. der Deut. Math.-Verein*, 3, 1892-/3, 107-156.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teubner, 1898 y 1924; Johnson Reprint Corp., 1965, vol. 3, 18-35, 748-829; vol. 4, 375-388.

Chasles, M.: *Aperçu historique sur l'origine et le développement des méthodes en géométrie* (1837), 3.ª edición, Gauthier-Villars, 1889, pp. 142-252.

Coolidge, Julian L.: «The beginnings of Analytic Geometry in Three Dimensions», *Amer. Math. Monthly*, 55, 1948, 76-86.

—: *A History of Geometrical Methods*, Dover (reimpresión), 1963, pp. 134-140 y 318-346.

—: *The Mathematics of Great Amateurs*. Dover (reimpresión, 1963, cap. 12.

—: *A History of Conic Sections and Quadric Surfaces*. Dover (reimpresión), 1968.

Euler, Leonhard: *Opera Omnia*, Orell Fussli, Serie 1, vol. 6 (1921), vols. 26-29 (1953-56); Serie 2, vols. 3 y 4 (1948-50), vol. 9 (1968).

Huygens, Christian: *Horologium Oscillatorium* (1673), reimpreso por Dawsons, 1966; también en Huygens, *Oeuvres complètes*, 18, 27-438.

Hofmann, Jos E.: Über Jakob Bernoullis Beiträge zur Infinitesimallmathematik», *L'enseignement mathématique*, (2), 2, 61-171, 1956, publicado por separado por el Instituto de Matemáticas, Ginebra, 1957.

Kotter, Ernst: «Die Entwickelung der synthetischen Geometrie von Monge bis auf Staudt», *Jahres. der Deut. Math.-Verein.*, 5, parte II, 1986, 1-486; también como libro, B. G. Teubner, 1901.

Lagrange, Joseph Louis: *Oeuvres*, Gauthier-Villars, 1867-69, vol. 1, 3-20; vol. 3, 619-692.

Lambert, J. H.: *Ammerkungen und Zusätze zur Entwefung der Land-und Himmelscharten* (1772), Ostwald's Klassiker núm. 54, Wielhelm Engelmann, Leipzig, 1896.

Loria, Gino: *Spezielle algebraische und tranzendente ebenen Kurven, Theorie und Geschichte*, 2 vols., 2.ª edición, B. G. Teubner, 1910-11.
Montucla, J. F.: *Histoire des Mathématiques* (1802), Albert Blanchard (reimpresión), 1960, vol. 3, pp. 63-102.
Struik, D. J.: *A Source Book in Mathematics (1200-1800)*, Harvard University Press, 1969, pp. 168-178, 180-183, 263-269, 413-419.
Taton, René: *L'oeuvre scientifique de Monge*, Presses Universitaires de France, 1951, cap. 4.
Whiteside, Derek T.: «Patters of Mathematical Thought in the Later Seventeenth Century», *Archive for History of Exact Sciences*, 1, 1961, 179-388. Véanse pp. 202-205 y 270-311.
—: *The Mathematical Works of Isaac Newton*, Johnson Reprint Corp., 1967, vol. 2, 137-161. Este contiene la *Enumeratio (Enumeración)* de Newton en inglés.

Capítulo 24
EL CALCULO DE VARIACIONES EN EL SIGLO XVIII

> Ya que la fábrica del universo es más que perfecta y es el trabajo de un Creador más que sabio, nada en el universo sucede en el que alguna regla de máximo o mínimo no aparezca.
>
> LEONHARD EULER

1. Los problemas iniciales

El cálculo de variaciones, durante sus etapas tempranas en el siglo que nos ocupa, difícilmente puede distinguirse del cálculo propiamente dicho, de modo análogo a lo que sucede en las áreas de series y ecuaciones diferenciales. No obstante, pocos años después de la muerte de Newton en 1727 se hacía evidente que una rama totalmente nueva de las matemáticas, con sus propios problemas y su metodología, había surgido. Esta nueva materia, casi comparable en importancia con las ecuaciones diferenciales para las matemáticas y las ciencias, suministró uno de los más grandes principios en toda la física matemática.

Para tener una noción preliminar de la naturaleza del cálculo de variaciones, consideremos los problemas que impulsaron a los matemáticos hacia este tema. Históricamente, el primer problema significativo fue propuesto y resuelto por Newton. En el libro segundo de sus *Principia* estudió el movimiento de los objetos en el agua; más adelante, en el escolio a la proposición 34 de la tercera edición, consideró el contorno que una superficie de revolución moviéndose a una velocidad constante en la dirección de sus ejes debe tener si

presenta la mínima resistencia al movimiento. Newton supuso que la resistencia del fluido en cualquier punto sobre la superficie del cuerpo es proporcional a la componente de la velocidad normal a la superficie. En los *Principia* proporcionó únicamente una caracterización geométrica del contorno deseado, pero en una misiva, presuntamente escrita a David Gregory en 1694, Newton dio su solución.

En forma moderna, el problema de Newton es encontrar el valor mínimo de la integral

$$J = \int_{x_1}^{x_2} \frac{y(x)[y'(x)]^3}{1 + [y'(x)]^2}\, dx$$

escogiendo la función adecuada $y(x)$ para el contorno de la curva que debe ser rotada alrededor del eje x (fig. 24.1). La característica peculiar de este problema (y de los problemas, en general, del cálculo de variaciones) es que propone una integral cuyo valor depende de una función desconocida $y(x)$ que aparece en el integrando y que debe ser determinada de tal manera que haga la integral máxima o mínima.

A pesar que Newton aplicó la idea de introducir un cambio en el contorno de la parte de meridiano $y(x)$, lo cual es casi igual a la esencia del método del cálculo de variaciones, su solución no es típica de la técnica del tema y por ello no la trataremos aquí. Puede ser de interés que las ecuaciones paramétricas de la $y(x)$ adecuada son

$$x = \frac{c}{p}(1 + p^2)^2, \qquad y = a + c\left(-\log p + p^2 + \frac{3}{4}\, p^4 \right),$$

donde p es el parámetro. Sobre este trabajo Newton dice: «Considero que esta proposición puede ser de utilidad en la construcción de barcos.» Los problemas de esta naturaleza han llegado a ser importantes no únicamente en el diseño de barcos, sino también de submarinos y de aeroplanos.

En el *Acta Eruditorum* de junio de 1696 [1], Jean Bernoulli propuso como un reto a otros matemáticos el ahora famosa problema de la braquistócrona. El problema es determinar la trayectoria descendente al resbalar una partícula en el tiempo mínimo desde un

[1] Página 269 = *Opera*, 1, 161.

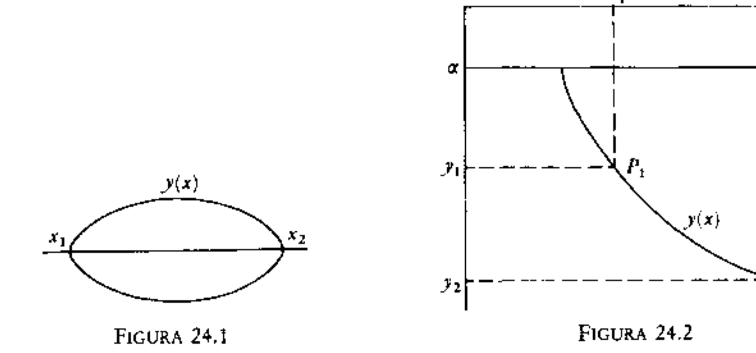


FIGURA 24.1 FIGURA 24.2

punto dado a otro punto no directamente debajo. La velocidad inicial v_1 en P_1 (fig. 24.2) es dada; y se descartan la fricción y la resistencia del aire. En forma moderna, este problema consiste en minimizar la integral J que representa el tiempo de descenso, donde

$$J = \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \sqrt{\frac{1 + [y'(x)]^2}{y(x) - \alpha}}\, dx.$$

Aquí g es la aceleración gravitacional y $\alpha = y_1 - v_1^2/2g$. De nuevo, la $y(x)$ en el integrando debe ser escogida de tal manera que haga a J tomar un mínimo. El problema había sido formulado y resuelto incorrectamente por Galileo (1630 y 1638), quien sugirió el arco de un círculo como respuesta. La respuesta correcta es el arco de la única cicloide que une P_1 y el segundo punto P_2, el cual es cóncavo hacia arriba; la línea l sobre la cual el círculo generador rueda debe estar justo a la altura adecuada, $y = \alpha$, por encima del punto inicial dado de caída. Entonces hay una y sólo una cicloide que pasa por los dos puntos.

Newton, Leibniz, L'Hospital, Jean Bernoulli y su hermano mayor Jacques encontraron la solución correcta: todas éstas fueron publicadas en el número de mayo de 1697 del *Acta Eruditorum*. Las soluciones de los dos Bernoulli merecen comentarios posteriores. El método de Jean [2] era ver que la trayectoria de descenso más rápido es la misma que la trayectoria de un rayo de luz en un medio con un índice de refracción adecuadamente seleccionado, $n(x, y) = c/\sqrt{y - \alpha}$. La ley

[2] *Acta Erud.*, 1697, 206-211 = *Opera*, 1, 187-193.

de la refracción en una discontinuidad brusca (ley de Snell) era conocida, de tal manera que Jean dividió el medio en un número finito de capas con un cambio súbito en el índice de capa a capa y más adelante hizo tender el número de capas al infinito. El método [3] de Jacques fue mucho más laborioso y geométrico, pero también más general y significó un paso más importante en dirección al método del cálculo de variaciones.

La cicloide era bien conocida a través del trabajo de Huygens y otros sobre el problema del péndulo (cap. 23, sec. 5). Cuando los hermanos Bernoulli encontraron que era también la solución para el problema de la braquistócrona, se sorprendieron. Jean Bernoulli dijo [4]: «Con justicia admiramos a Huygens porque fue el primero en descubrir que una partícula pesada recorre una cicloide en el mismo tiempo, sin importar cuál sea el punto de partida. Pero usted quedará sorprendido cuando yo diga que esta misma cicloide, la tautócrona de Huygens, es la braquistócrona que hemos estado buscando.»

Otra clase importante de problemas es el de las geodésicas, esto es, las trayectorias de longitud mínima entre dos puntos sobre una superficie. Si la superficie es un plano, entonces la integral es

$$J = \int_{x_1}^{x_2} \sqrt{1 + [y'(x)]^2}\, dx,$$

y, por supuesto, la respuesta es un segmento lineal. En el siglo XVIII, el problema geodésico de mayor interés estaba relacionado con las trayectorias mínimas sobre la superficie de la Tierra, cuyo contorno preciso no era conocido, aunque los matemáticos creían que se trataba de alguna forma de elipsoide y más posiblemente de una figura de revolución. El trabajo pionero sobre geodésicas ya mencionado (cap. 23, sec. 7) no empleó los métodos del cálculo de variaciones, pero era claro que los trucos especiales no serían lo suficientemente poderosos como para tratar el problema general de las geodésicas.

Analíticamente, los problemas hasta entonces formulados son de la forma

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx$$

---

[3] *Acta Erud.*, 1697, 211-217 = *Opera*, 2, 768-778.

[4] *Opera*, 1, 187-193.

y requieren encontrar la $y(x)$ que se extiende de (x_1, y_1) a (x_2, y_2) y que minimiza o maximiza J. Otra clase de problemas, llamados problemas isoperimétricos, entraron también en la historia del cálculo de variaciones al final del siglo XVII. El precursor de esta clase de problemas —de todas las curvas planas cerradas con un perímetro dado, encontrar la que encierra un área máxima— puede datar de antes de los tiempos de los griegos. Existe una historia referente a que la princesa Dido de los antiguos fenicios de la ciudad de Tiro huyó de su casa para establecerse sobre la costa mediterránea del norte de Africa. Allí compró alguna tierra y accedió a pagar una suma fija por tanta tierra como podría ser rodeada por la piel de un toro. La sutil Dido cortó la piel en tiras muy delgadas, ató las tiras una con otra y procedió a encerrar una área que tenía la longitud total de estas tiras como perímetro. Más aún, escogió tierra junto al mar, de tal manera, que no se necesitara piel a lo largo de la costa. De acuerdo con la leyenda, Dido decidió que la longitud de la piel debería formar un semicírculo —la forma correcta que encierra una área máxima.

Aparte del trabajo de Zenodoro (cap. 5, sec. 7), no hubo prácticamente trabajos sobre problemas isoperimétricos hasta el final del siglo XVII. En un intento por retar y apenar a su hermano, Jacques Bernoulli propuso un problema isoperimétrico bastante complicado, que contenía varios casos, en el *Acta Eruditorum* de mayo de 1697 [5]. Jacques aún ofreció a Jean un premio de cincuenta ducados por una solución satisfactoria. Jean dio varias soluciones, una de las cuales fue obtenida en 1701 [6], pero todas eran incorrectas. Jacques proporcionó una solución correcta [7], y los hermanos pelearon sobre la veracidad de las soluciones del otro. De hecho, el método de Jacques, como en el caso del problema de la braquistócrona, fue un gran paso hacia la técnica general que pronto estaría en boga. En 1718, Jean [8] mejoró notablemente la solución de su hermano.

Analíticamente, el problema isoperimétrico básico está formulado así. Las curvas posibles se representan paramétricamente por

$$x = x(t), \qquad y = y(t), \qquad t_1 \leq t \leq t_2$$

---

[5] Página 214.
[6] *Mém. de l'Acad. des Sci., Paris*, 1706, 235 = *Opera*, 1, 424.
[7] *Acta Erud.*, 1701. 213 sigs. = *Opera*, 2, 897-920.
[8] *Mém. de l'Acad. des Sci., Paris*, 1718, 100 sigs. = *Opera*, 2, 235-269.

y, ya que son curvas cerradas, $x(t_1) = x(t_2)$ e $y(t_1) = y(t_2)$. Más aún, ninguna curva se debe cortar a sí misma. Más adelante, el problema requiere determinar las $x(t)$ e $y(t)$ tales que la longitud

$$L = \int_{t_1}^{t_2} \sqrt{(x')^2 + (y')^2}\, dt$$

es una constante dada y tal que la integral de área

$$J = \int_{t_1}^{t_2} (xy' - x'y)\, dt$$

tiene un máximo. Hay dos características notables en este problema isoperimétrico. Uno, el uso de la representación paramétrica es incidental. El otro es la presencia de una condición auxiliar, que L debe ser constante.

Determinar el contorno de la curva a lo largo del cual una partícula se desliza desde un punto dado P_1, con una velocidad inicial dada v_1, a cualquier punto de la línea l (fig. 24.3) de tal manera que el tiempo de deslizamiento de P_1 a l sea mínimo, fue otro problema que Jacques propuso en el número de mayo de 1697 del *Acta.* Aquí difiere de los anteriores en que las curvas posibles no se extienden de un punto fijo a otro, sino de algún punto fijo a alguna línea. La respuesta, proporcionada por Jacques en el *Acta* de 1698 (aunque en posesión de Jean en 1697, inédita), es un arco de una cicloide que corta a la línea l en ángulos rectos. Este problema fue más tarde generalizado a los casos donde l puede ser cualquier curva dada y donde, en lugar de P_1, se da otra curva de tal manera que el problema es encontrar la trayectoria que necesita el menor tiempo para deslizarse desde algún punto sobre una curva dada a algún punto

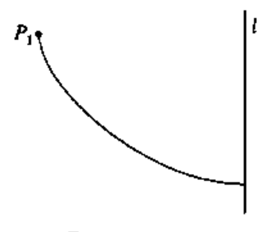


FIGURA 24.3

sobre otra. Esta clase de problemas se encierra en la frase «problemas con extremos variables».

2. Los primeros trabajos de Euler

En 1728, Jean Bernoulli propuso a Euler el problema de obtener geodésicas sobre superficies aplicando la propiedad que los planos osculadores de las geodésicas cortan la superficie en ángulos rectos (cap. 23, sec. 7). Este problema inició a Euler en el cálculo de variaciones. Lo resolvió en 1728 [9]. En 1734 Euler generalizó el problema de la braquistócrona para minimizar cantidades diferentes del tiempo, y tomando en cuenta un medio resistente [10].

Más adelante, Euler se propuso encontrar una aproximación más general a problemas en este terreno. Su método, que fue una simplificación del de Jacques Bernoulli, consistió en reemplazar la integral de un problema por una suma y reemplazar las derivadas en el integrando por cocientes diferenciales, haciendo de la integral una función de un número finito de ordenadas del arco $y(x)$. Más adelante, varió una o más de las ordenadas seleccionadas arbitrariamente y calculó la variación en la integral. Igualando la variación de la integral a cero y usando un proceso de paso al límite muy tosco para transformar las ecuaciones en diferencias resultantes, obtuvo la ecuación diferencial que debía ser satisfecha por el arco minimizante.

Por el método descrito con anterioridad, aplicado a integrales de la forma

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx, \qquad (1)$$

Euler tuvo éxito al demostrar que la función $y(x)$ que minimiza o maximiza el valor de J debe satisfacer la ecuación diferencial ordinaria

$$f_y - \frac{d}{dx}(f_{y'}) = 0. \qquad (2)$$

---

[9] *Comm. Acad. Sci. Petrop.*, 3, 1728, 110-124, pub. 1732 = *Opera*, (1), 25, 1-12.

[10] *Comm. Acad. Sci. Petrop.*, 7, 1734/35, 135-149, pub. 1740 = *Opera*, (1), 25, 41-53.

Esta notación ha de ser entendida de la siguiente manera: el integrando $f(x, y, y')$ debe ser visto como una función de las variables independientes x, y e y' en cuanto concierne a f_y y $f_{y'}$. Sin embargo, $df_{y'}/dx$ se toma como la derivada de $f_{y'}$, donde $f_{y'}$ depende de x a través de x, y e y'. Esto es, la ecuación diferencial de Euler es equivalente a

$$f_y - f_{y'x} - f_{y'y}y' - f_{y'y'}y'' = 0. \qquad (3)$$

Ya que f es conocida, esta es una ecuación diferencial ordinaria de segundo orden, generalmente no lineal, en $y(x)$. Esta famosa ecuación, que Euler publicó en 1736 [11], es aún la ecuación diferencial básica del cálculo de variaciones. Es, como veremos con mayor claridad más adelante, una condición necesaria que la función $y(x)$ minimizante o maximizante debe satisfacer.

Después Euler atacó problemas más difíciles con condiciones de contorno especiales, como los problemas isoperimétricos, pero su procedimiento seguía siendo resolver la ecuación diferencial (3) para primero obtener los posibles arcos minimizantes o maximizantes y entonces determinar a partir del número de constantes en las soluciones generales de (2) o (3) qué condiciones adicionales podía aplicar. Daniel Bernoulli, en una carta de 1742, fue quien llamó su atención sobre uno de los problemas que atacó. Bernoulli propuso encontrar el contorno de una cuerda elástica sujeta a presión en ambos extremos al suponer que el cuadrado de la curvatura a lo largo de la curva con la varilla doblada, esto es, $\int_0^L ds/R^2$, donde s es el arco de longitud y R, el radio de la curvatura, es un mínimo. Esta condición implica suponer que la energía potencial almacenada en la forma tomada por la varilla es mínima.

La ecuación diferencial (3) no es la adecuada cuando los integrandos de las integrales minimizados o maximizados son más complicados que (1). En los años de 1736 a 1744, Euler mejoró sus métodos y obtuvo ecuaciones diferenciales análogas a (3) para un buen número de problemas. Estos resultados los publicó en un libro de 1744, *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes (El arte de encontrar líneas curvas que gozan de algunas propiedades máximas o mínimas)* [12]. El trabajo de Euler en

[11] *Comm. Acad. Sci. Petrop.*, 8, 1736, 159-190, pub. 1741 = *Opera*, (1), 25, 54-80.
[12] *Opera*, (1), 24.

su *Methodus* fue pesado porque utilizó consideraciones geométricas, diferencias sucesivas y series, y cambió derivadas por cocientes diferenciales e integrales en sumas finitas; en otras palabras, fracasó en hacer más eficaz el uso del cálculo. Pero acabó obteniendo fórmulas simples y elegantes aplicables a una gran variedad de problemas, empleando gran cantidad de ejemplos para mostrar la conveniencia y generalidad de su método. Un ejemplo se refiere a superficies mínimas de revolución. Aquí el problema es determinar la curva plana $y = f(x)$ entre (x_0, y_0) y (x_1, y_1), de tal manera que cuando gira alrededor del eje x genera la superficie de área mínima. La integral que ha de ser minimizada es

$$A = \int_{x_0}^{x_1} 2\pi y \sqrt{1 + y'^2}\, dx. \tag{4}$$

Euler probó que la función $f(x)$ debe ser el arco de catenaria; la superficie así generada es llamada catenoide. En un apéndice a su libro de 1744, Euler también proporcionó una solución definitiva al problema de la varilla elástica mencionado con anterioridad. No solamente dedujo que el contorno de la vara debía tomar la forma dada por una integral elíptica, sino que también dio soluciones para diferentes tipos de condiciones. Este libro le trajo fama inmediata y reconocimiento como el más grande de los matemáticos vivos.

Con dicho trabajo, el cálculo de variaciones llegó a su existencia como una nueva rama de las matemáticas. Sin embargo, los argumentos geométricos fueron aplicados extensamente, y los argumentos analíticos y geométricos combinados no solamente resultaban complicados, sino que difícilmente proporcionaron un método general sistemático. Euler era completamente consciente de estas limitaciones.

3. El principio de mínima acción

Mientras que se progresaba en la solución de los problemas del cálculo de variaciones, una nueva motivación para trabajar en la materia llegó directamente de la física: este desarrollo contemporáneo fue el principio de acción mínima. Para explicar las bases de este principio debemos retroceder un poco: Euclides había demostrado en la *Catoptrica* (cap. 7, sec. 7) que la luz que viaja de P a un espejo

(fig. 24.4) y después a Q sigue la trayectoria para la cual $\sphericalangle 1 = \sphericalangle 2$. Entonces, el alejandrino Herón demostró que la trayectoria PRQ, la que de hecho sigue la luz, es más corta que cualquier otra trayectoria, tal como $PR'Q$, que pudiera concebirse que tomara. Ya que la luz sigue la trayectoria más corta, si el medio en el lado superior de la línea RR' es homogéneo, entonces la luz viaja con velocidad constante y por lo tanto sigue la trayectoria que requiere el tiempo mínimo. Herón aplicó este principio de trayectoria y tiempo mínimos a problemas de reflexión, a partir de espejos esféricos cóncavos y convexos.

Basando sus asertos sobre este fenómeno de reflexión y en principios filosóficos, teológicos y estéticos, filósofos y científicos posteriores a los griegos propusieron la doctrina que la naturaleza actúa de la manera mínima, como Olimpiodoro (s. VI d.C.) dijo en su *Catoptrica*: «La naturaleza no hace algo superfluo ni cualquier trabajo innecesario.» Leonardo da Vinci afirmó que la naturaleza es económica y que su economía es cuantitativa, y Roberto Grosseteste creía que la naturaleza siempre actúa en la mejor, mínima y matemáticamente manera posible. En los tiempos medievales era aceptado comúnmente que la naturaleza se comportaba de esta manera.

Los científicos del siglo XVII fueron al menos receptivos a esta idea, pero, como científicos, intentaron acercarlo al fenómeno que lo apoyaba. Fermat sabía que bajo la reflexión la luz sigue la trayectoria que requiere el mínimo tiempo y, convencido de que la naturaleza actúa simple y económicamente, afirmó en misivas de 1657 y 1662 [13] su principio de tiempo mínimo, el cual afirma que la luz siempre sigue la trayectoria que requiere el mínimo tiempo. Había dudado de la validez de la ley de la refracción de la luz (cap. 15,

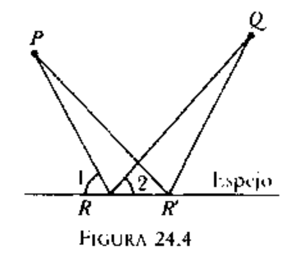


FIGURA 24.4

[13] *Œuvres*, 2, 354-359 y 457-463.

sec. 4), pero cuando encontró en 1661 [14] que la podía deducir de su propio principio, no solamente resolvió sus dudas acerca de la ley, sino que sintió aún con mayor certeza que su principio era correcto.

El principio de Fermat se puede expresar matemáticamente en varias formas equivalentes. De acuerdo con la ley de refracción

$$\frac{\operatorname{sen} i}{\operatorname{sen} r} = \frac{v_1}{v_2},$$

donde v_1 es la velocidad de la luz en el primer medio y v_2 en el segundo. La proporción de v_1 a v_2 se denota por n y es llamado el índice de refracción del segundo medio relativo al primero, o, si el primero es el vacío, n es llamado índice absoluto de refracción del medio no vacío. Si c denota la velocidad de la luz en el vacío, entonces el índice absoluto es $n = c/v$, donde v es la velocidad de la luz en el medio. Si el medio es variable de punto a punto, entonces n y m son funciones de x, y y z. De aquí que el tiempo requerido por la luz para viajar del punto P_1 al punto P_2 a lo largo de la curva $x(\sigma)$, $y(\sigma)$, $z(\sigma)$ esté dado por

$$J = \int_{\sigma_1}^{\sigma_2} \frac{ds}{v} = \int_{\sigma_1}^{\sigma_2} \frac{n}{c}\, ds$$

$$= \frac{1}{c} \int_{\sigma_1}^{\sigma_2} n(x, y, z) \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, d\sigma, \qquad (5)$$

donde σ_1 es el valor de σ en P_1 y σ_2 el valor de P_2. Así, el principio establece que la trayectoria que la luz toma de hecho al viajar de P_1 a P_2 está dada por la curva que hace J mínimo [15].

Al principio del siglo XVIII los matemáticos tenían varios ejemplos impresionantes del hecho de que la naturaleza intenta maximizar o minimizar ciertas cantidades importantes. Huygens, quien en un principio había objetado al principio de Fermat, demostró que funciona para la propagación de la luz en medios con índices variables de refracción. Incluso la primera ley del movimiento de Newton, que la línea recta o la distancia mínima es el movimiento natural

[14] *Œuvres*, 2, 457-463.

[15] Existen casos, como por ejemplo en la reflexión de la luz de un espejo cóncavo, donde la luz toma la trayectoria que requiere el tiempo máximo. Este hecho era conocido de Fermat y fue establecido explícitamente por William R. Hamilton.

de un cuerpo, mostró el deseo de economizar de la naturaleza. Estos ejemplos sugerían que podía haber algún principio más general. La búsqueda de tal principio fue realizada por Maupertuis.

Pierre-Louis Moreau de Maupertuis (1698-1759), mientras trabajaba en la teoría de la luz en 1744, propuso su famoso principio de acción mínima en su ensayo titulado «Accord des différentes lois de la nature qui avaient jusqu'ici paru incompatibles» («Acuerdo de las diferentes leyes de la naturaleza que hasta ahora parecían incompatibles»)[16]. Empezó por el principio de Fermat, pero debido a las diferencias en aquel tiempo sobre si la velocidad de la luz era proporcional al índice de refracción como creían Descartes y Newton, o inversamente proporcional como creía Fermat, Maupertuis abandonó el tiempo mínimo. De hecho, no creía que siempre fuera correcto.

La acción, decía Maupertuis, es la integral del producto de la masa, velocidad y distancia recorrida y cualesquiera cambios en la naturaleza son tales que hacen la acción mínima. Maupertuis fue algo vago, ya que fracasó en especificar el intervalo de tiempo sobre el cual el producto de m, v y s debía ser tomado y porque asignaba un significado diferente a la acción en cada una de las aplicaciones que hacía a la óptica y en algunos problemas de mecánica.

A pesar de que tenía algunos ejemplos físicos en los cuales apoyar su principio, Maupertuis lo defendía también por razones teológicas. Las leyes del comportamiento de la materia tenían que poseer la perfección que merecía la creación de Dios; y el principio de acción mínima parecía satisfacer este criterio porque mostraba que la naturaleza era económica. Maupertuis proclamó que su principio era una ley universal de la naturaleza y la primera prueba científica de la existencia de Dios. Euler, quien entre 1740 y 1744 había mantenido correspondencia con Maupertuis sobre este tema, coincidía con éste en que Dios debía haber construido el universo de acuerdo con algún principio básico y que la existencia de dicho principio evidenciaba la mano de Dios.

En el segundo apéndice de su libro de 1744, Euler formuló el principio de acción mínima como un teorema dinámico exacto. Se limitó al movimiento de una partícula individual moviéndose a lo largo de curvas planas. Más aún, suponía que la velocidad dependía

---

[16] *Mém. de l'Acad. des Sci., Paris*, 1744.

de la posición o, en términos modernos, que la fuerza es derivable de un potencial. Mientras que Maupertuis escribía

$$mvs = \text{mín.},$$

Euler escribió

$$\partial \int v \; ds = 0,$$

por lo que entendía que la proporción de cambio de la integral por el cambio en la trayectoria debe ser cero. También escribió que, ya que $ds = \; dt$, debe ser

$$\partial \int v^2 \; ts = 0.$$

Que entendía Euler por la proporción de cambio de la integral era aún muy vago aquí, a pesar de que aplicó correctamente el principio en problemas específicos mediante el uso de su técnica del cálculo de variaciones. Al menos mostró que la acción de Maupertuis era mínima para movimientos a lo largo de curvas planas.

Euler fue más lejos que Maupertuis al creer que todos los fenómenos naturales se comportaban buscando maximizar o minimizar alguna función, de tal manera que los principios físicos básicos debían ser expresados como consecuencia de que alguna función sea maximizada o minimizada. En particular, esto debía ser cierto en la dinámica, que estudia los movimientos de los cuerpos propulsados por fuerzas. Euler no estaba muy lejos de la verdad.

4. La metodología de Lagrange

Lagrange, cuando sólo contaba diecinueve años, empezó a interesarse con Euler por problemas del cálculo de variaciones. Descartó los argumentos geometrico-analíticos de los Bernoulli y Euler e introdujo métodos puramente analíticos. En 1755 obtuvo un procedimiento general, sistemático y uniforme para una gran variedad de problemas, y trabajó en ellos por muchos años. Su más famosa publicación en esta materia fue su «Essai d'une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfi-

nies» («Ensayo sobre un nuevo método para determinar los máximos y los mínimos de fórmulas integrales indefinidas») [17]. En una carta a Euler de agosto de 1755, describió su método, al que llamó el método de variaciones, pero que Euler presentó en un ensayo a la Academia de Berlín en 1756 [18] llamándolo cálculo de variaciones.

Veamos el método de Lagrange para el problema básico del cálculo de variaciones, es decir, minimizar o maximizar la integral

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx, \tag{6}$$

donde $y(x)$ debe ser determinada. Una de las innovaciones de Lagrange consistió en no variar las ordenadas de la curva $y(x)$ minimizadora o maximizadora, sino en introducir nuevas curvas entre los puntos extremos (x_1, y_1) y (x_2, y_2). Estas nuevas curvas (fig. 24.5) fueron representadas por Lagrange en la forma $y(x) + \delta y(x)$, siendo δ un símbolo especial introducido por Lagrange para indicar la variación de la *curva entera* $y(x)$. La introducción de una nueva curva en el integrando de (6) cambia desde luego el valor de J. El incremento en J, el cual denotamos por ΔJ, es entonces

$$\Delta J = \int_{x_1}^{x_2} \{f(x, y + \delta y, y' + \delta y') - f(x, y, y')\}\, dx.$$

Ahora bien, Lagrange considera f como una función de tres variables independientes, pero debido a que x no cambia, el integrando puede ser escrito por medio del teorema de Taylor aplicado a una función

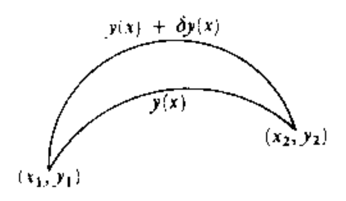


FIGURA 24.5

[17] *Misc. Taur.*, 2, 1760/61, 173-195, pub. 1762 = *Œuvres*, 1, 333-362.

[18] «Elementa Calculi Variationum» («Elementos del cálculo de variaciones»), *Novi Comm. Acad. Sci. Petrop.*, 10, 1764, 51-93, pub. 1766 = *Opera*, (1), 25, 141-176.

de dos variables. La expansión produce términos de primer orden en δy y en $\delta y'$, términos de segundo grado en estos incrementos, y así sucesivamente. Entonces, Lagrange escribe

$$\Delta J = \delta J + \frac{1}{2}\ \delta^2 J + \ \frac{1}{3!}\,\delta^3 J + \cdots \tag{7}$$

donde δJ indica la integral de los términos de primer grado en δy y en $\delta y'$, $\delta^2 J$ indica la integral de los términos de segundo grado, y, así de aquí en adelante. Así que

$$\delta J = \int_{x_1}^{x_2} (f_y \delta y + f_{y'} \delta y')\, dx$$

$$\delta^2 J = \int_{x_1}^{x_2} \{f_{yy}\, (\delta y)^2 + 2 f_{yy'} (\delta y)\, (\delta y') + f_{y'y'}\, (\delta y')^2\}\, dx.$$

δJ es llamado primera variación de J; $\delta^2 J$, la segunda variación y las demás análogamente.

Lagrange argumenta ahora que el valor de δJ, ya que contiene los términos de primer orden en las pequeñas variaciones δy y $\delta y'$, domina en el lado derecho de (7), de tal manera que cuando δJ es positivo o negativo, ΔJ será positivo o negativo. Pero en el máximo o mínimo de J, ΔJ debe tener el mismo signo, como en el caso de máximos y mínimos ordinarios de una función $f(x)$ de una variable, de tal forma que para que $y(x)$ sea una función maximizante, δJ debe ser 0. Más aún, Lagrange dice que

$$\delta y' = \frac{d(\delta y)}{dx}\,; \tag{8}$$

esto es, el orden de las operaciones d y δ puede ser cambiado. Esto es correcto, aunque la razón no estaba clara para los contemporáneos de Lagrange, y Euler la clarificó más tarde. (Es muy fácil que se vea correctamente si escribimos $y + \delta y$ como $y + n(x)$, donde $n(x)$ es la variación de $y(x)$. Entonces $\delta y = y \ + n(x) - y = n(x)$ y $\delta y' = y' + n'(x) - y' = n'(x)$. Pero $n'(x) = dn(x)/dx = d(\delta y)/dx$.) Usando (8), Lagrange escribe la primera variación como

$$\delta J = \int_{x_1}^{x_2} \left[\ f_y\ \delta y\ + f_{y'}\ \frac{d}{dx}\ (\delta y)\ \right] dx.$$

Integrando el segundo término por partes y usando el hecho de que δy debe anularse en x_1 y en x_2, se obtiene

$$\delta J = \int_{x_1}^{x_2} \left(f_y \delta y - \left(\frac{d}{dx} f_{y'} \right) \delta y \right) dx. \tag{9}$$

Ahora δJ debe ser 0 para toda variación δy. De aquí Lagrange concluye que el coeficiente de δy debe ser 0, [19] o que

$$f_y - \frac{d}{dx}(f_{y'}) = 0. \tag{10}$$

Así llegó Lagrange a la misma ecuación diferencial ordinaria para $y(x)$ que había obtenido Euler. El método de Lagrange de derivar (10) (excepto por su uso de diferenciales), y aun su notación, son usados hoy en día. Por supuesto, (10) es una condición necesaria para $y(x)$ pero no suficiente.

En este ensayo de 1760-61, Lagrange dedujo también, por primera vez, condiciones finales que deben ser satisfechas por una curva minimizadora para problemas con puntos finales variables. Encontró las condiciones de transversalidad que deben cumplirse en las intersecciones de la curva minimizadora con las curvas fijas, o superficies, sobre las cuales los puntos finales de las curvas de comparación pueden variar (fig. 24.6).

Aunque queda mucho por decir acerca de cómo maximizar o minimizar integrales de la forma (6), históricamente el siguiente paso, dado por Lagrange en su ensayo de 1760-61 y en el siguiente [20], fue considerar problemas que lo condujeron a integrales múltiples. La integral que ha de ser maximizada o minimizada es de la forma

$$J = \iint f(x, y, z, p, q)\, dx\, dy, \tag{11}$$

---

[19] El hecho de que el coeficiente de δy debe ser cero era aceptado intuitivamente o demostrado incorrectamente por todo autor sobre la materia, cien años después del trabajo de Lagrange. Incluso la demostración de Cauchy era inadecuada. La primera prueba correcta fue proporcionada por Pierre Frédéric Sarrus (1798-1861) (*Mém. divers Savans*, [2], 10, 1848, 1-128). El resultado es hoy conocido como el lema fundamental del cálculo de variaciones.

[20] *Misc. Taur.*, 4, 1766/69 = *Œuvres*, 2, 37-63.

donde z es una función de x e y, $p = \partial z/\partial x$ y $q = \partial z/\partial y$. La integración es sobre algún área en el plano xy. El problema, entonces, es encontrar la función $z(x, y)$ que maximiza o minimiza el valor de J. Uno de los problemas más importantes que se encuentra con esta clase de integrales dobles es hallar la superficie de área mínima entre todas las superficies cuyas fronteras están fijas de alguna manera. Así pueden darse dos curvas cerradas que no se intersecan en el espacio y buscar la superficie de área mínima encerrada por estas dos curvas. Como un caso especial del problema de la superficie mínima, las dos curvas pueden ser círculos paralelos al plano yz (fig. 24.7) y con centros sobre el eje x. Entonces las posibles superficies mínimas son necesariamente superficies de revolución de área mínima. Este último problema, como señalamos con anterioridad, ya había sido resuelto por Euler en 1744. Sin embargo, es posible tratar el caso especial de la superficie de revolución con la teoría aplicable a la integral (11).

Por un método similar al que había usado para la integral simple (6), Lagrange obtuvo la ecuación diferencial que la función $z(x, y)$ minimizadora de (11) debía satisfacer. Si usamos la notación habitual

$$\frac{\partial z}{\partial x} = p, \qquad \frac{\partial z}{\partial y} = q, \qquad \frac{\partial^2 z}{\partial x^2} = r, \qquad \frac{\partial^2 z}{\partial x \partial y} = s, \qquad \frac{\partial^2 z}{\partial y^2} = t,$$

entonces la ecuación es

$$Rr + Sc + Tt = U, \tag{12}$$

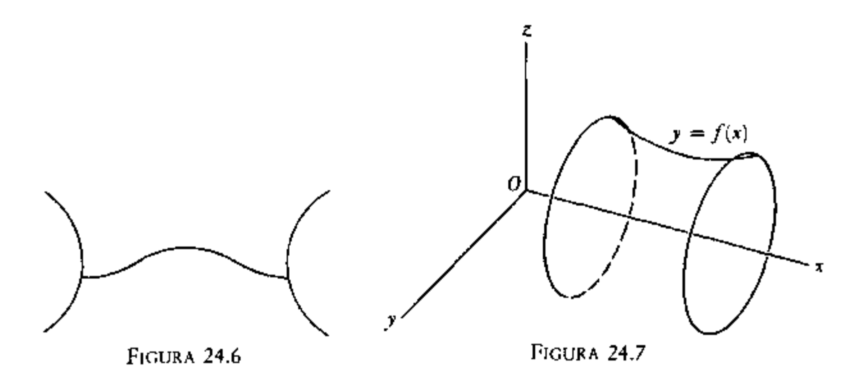


FIGURA 24.6 FIGURA 24.7

donde R, S, T y U son funciones de x, y, z, p y q. Esta ecuación diferencial parcial de segundo orden no lineal, llamada ecuación de Monge, no es fácil de resolver; ecuaciones de este tipo han sido materia de investigación desde la época de Euler hasta hoy (cap. 22, sec. 7).

En el caso del problema de la superficie mínima, la integral (1) se convierte en

$$\iint (1 + p^2 + q^2)^{1/2}\, dx\, dy, \tag{13}$$

y para esta clase especial de problemas la ecuación diferencial parcial (12) se convierte en

$$(1 + q^2)r - 2pqs + (1 + p^2)t = 0 \tag{14}$$

La ecuación antes expresada está en el ensayo de Lagrange, publicado en 1760-61 (aunque no precisamente en esta forma) y es un resultado analítico fundamental en la teoría de las superficies mínimas. Geométricamente, como señaló Meusnier en su ensayo de 1785 [21], dicha ecuación diferencial parcial expresa el hecho que sobre cualquier punto de la superficie minimizante los radios principales de curvatura son iguales y opuestos, o que la curvatura media, esto es, el promedio de las curvaturas principales, es cero.

Lagrange, en un ensayo posterior (1770) [22], también consideró integrales simples y múltiples en las cuales aparecen en el integrando derivadas superiores a las primeras. Este tema ha sido bien desarrollado desde los tiempos de Lagrange y es hoy material común en el cálculo de variaciones. Sin embargo, ya que los principios no son básicamente diferentes de los casos considerados con anterioridad, no nos iremos hacia este tema. En su *Mécanique analytique (Mecánica analítica)* se incorporan los contenidos de los ensayos sobre el cálculo de variaciones.

El cálculo de variaciones no fue muy bien entendido por los contemporáneos de Lagrange y Euler. Este explicó el método de Lagrange en numerosos escritos y lo usó para probar de nuevo un buen número de viejos resultados. A pesar de haberse percatado de

---

[21] *Mém. divers Savans*, 10, 1785, 477-485.

[22] *Nouv. Mém. de l'Acad. de Berlin*, 1770 = *Œuvres*, 3, 157-186.

que el cálculo de variaciones era una nueva rama o técnica, de la cual dice que está representada por el nuevo símbolo operacional δ, Euler, al igual que Lagrange, intentó basar la lógica del cálculo de variaciones en el cálculo ordinario. La idea de Euler [23] fue introducir un parámetro t tal que las curvas de la familia consideradas en un problema de variaciones variaría con t, esto es, para cada t en algún rango habría una curva $y(x)$. Entonces, dice Euler, mientras que $dy = (dy/dx)\,dx$, $\delta y = (dy/dt)dt$. De aquí que la variación δ sea expresable por una derivada parcial con respecto a t. Más adelante formuló la técnica del cálculo de variaciones en términos de su nuevo concepto de diferenciación con respecto a t. Sus resultados finales fueron, por supuesto, los mismos que ya había obtenido.

Euler, durante 1779 [24], consideró curvas espaciales con propiedades de máximo y mínimo y, en 1780, extensiones del problema de la baquistócrona cuando la fuerza aplicada (la cual es la gravedad en el problema usual) opera en tres dimensiones, o cuando está presente un medio resistente [25].

5. Lagrange y la mínima acción

Lagrange aplicó el cálculo de variaciones a la dinámica. Tomó de Euler el principio de la acción mínima, siendo el primero en expresar el principio en forma correcta, es decir, que para una partícula individual la integral del producto de la masa, velocidad y distancia tomada entre dos puntos fijos es un máximo o un mínimo; esto es, $\int mv\, ds$ debe ser un máximo o un mínimo para la trayectoria que de hecho sigue la partícula. De forma alternativa, ya que $ds = v\, dt$, entonces $\int mv^2\, dt$ debe ser máximo o mínimo. La cantidad mv^2 [hoy en día $(1/2)\, mv^2$] es llamada energía cinética; en los días de Lagrange era llamada fuerza viva. Lagrange también afirmó que el principio es cierto para una colección de partículas y aun para masas extensas, aunque no tenía mucha seguridad en esto último.

Usando el principio de mínima acción y el método del cálculo

---

[23] *Novi Comm. Acad. Sci. Petrop.*, 16, 1771, 35-70, pub. 1772 = *Opera*, (1), 25, 208-235.

[24] *Mém. de l'Acad. des Sci. de St. Peters.*, 4, 1811, 18-42, pub. 1813 = *Opera*, (1), 25, 293-313.

[25] *Mém. de l'Acad. des Sci. de St. Peters.*, 8, 1817/18, 17-45, pub. 1822 = *Opera*, (1), 25, 314-342.

de variaciones, Lagrange obtuvo sus famosas ecuaciones del movimiento. Consideremos el caso donde la energía cinética es una función de x, y y z. Entonces para una partícula individual la energía cinética T está dada por

$$T = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2). \tag{15}$$

Lagrange también supuso que las fuerzas que causan el movimiento eran derivables de la función potencial V, la cual depende de x, y y z. Una condición adicional, entonces, es que $T + V =$ const., esto es, la energía total es constante. La acción de Lagrange es

$$\int_{t_0}^{t_1} T\,dt \tag{16}$$

y su principio de acción mínima establece que la acción debe ser un mínimo o un máximo, esto es, que

$$\delta \int_{t_0}^{t_1} T\,dt = 0. \tag{17}$$

En una acción minimizante o maximizante, aun si el movimiento tiene lugar entre dos puntos fijos en el espacio y los dos valores de tiempo fijos t_0 y t_1, las variables de tiempo y espacio deben variar.

Aplicando el método del cálculo de variaciones a la integral de acción, Lagrange derivó ecuaciones análogas a las ecuaciones (2) de Euler, a saber,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial x}\right) + \frac{\partial V}{\partial x} = 0 \tag{18}$$

y las dos ecuaciones correspondientes con y y z. Estas ecuaciones son equivalentes a la segunda ecuación del movimiento de Newton.

Lagrange realizó el paso adelante siguiente de introducir las que ahora son llamadas coordenadas generalizadas. Esto es, en lugar de coordenadas rectangulares uno puede usar coordenadas polares o, de hecho, cualquier conjunto de coordenadas q_1, q_2, q_3, que son necesarias para fijar la posición de la partícula (o masa extensa). Entonces

$$x = x(q_1, q_2, q_3)$$
$$y = y(q_1, q_2, q_3)$$
$$z = z(q_1, q_2, q_3),$$

donde las q_i son ahora funciones de t. En términos de las nuevas coordenadas, T se convierte en una función de q_i y $\dot{q}_i$ mientras que V es una función de las q_i. Entonces las ecuaciones (18) se convierten en

$$\frac{d}{dt}\left(\frac{\partial T}{\partial x}\right) - \frac{\partial T}{\partial \dot{q}_i} + \frac{\partial V}{\partial q_i} = 0, \quad i = 1, 2, 3. \tag{19}$$

Este es un conjunto de tres ecuaciones diferenciales ordinarias de segundo orden simultáneas en las $\dot{q}_i$. Estas son las ecuaciones (características) de Euler para la integral de acción. Si n coordenadas son necesarias para fijar la posición del objeto en movimiento (por ejemplo, dos partículas requieren seis coordenadas), entonces las ecuaciones (19) son reemplazadas por n ecuaciones [26].

Estas coordenadas generalizadas no tienen necesariamente un significado geométrico o físico. Hoy en día, uno habla de ellas como coordenadas de un espacio de configuración y entonces las $q_i(t)$ son las ecuaciones de una trayectoria en el espacio de configuración. Así, Lagrange reconoció que el principio variacional, esto es, que la acción debe ser un mínimo o un máximo, puede ser usado con cualquier conjunto de coordenadas y las ecuaciones lagrangianas del movimiento (19) son invariantes en su forma con respecto a cualquier transformación de coordenadas.

El principio de Lagrange, no obstante equivaler a la segunda ley del movimiento de Newton, tiene varias ventajas sobre la formulación de Newton. Ante todo, cualquier sistema coordenado conveniente ya está, por así decirlo, construido en la formulación. Segundo, es más fácil manejar problemas con restricciones en el movimiento. Tercero, en lugar de una serie de ecuaciones diferenciales separadas, que podrían ser muy numerosas si se consideran muchas partículas, hay —para empezar, al menos— un principio del cual se siguen las ecuaciones diferenciales. Finalmente, a pesar de que su

---

[26] Lagrange es explícito al decir que el número de variables en T y en V es el número requerido para determinar la posición del sistema mecánico. Así, si hay N partículas independientes y cada una requiere tres coordenadas (x_i, y_i, z_i) para describir su trayectoria en el espacio, entonces las $3N$ coordenadas son necesarias. En este caso habrá $3N$ coordenadas q_i, $3N$ ecuaciones relacionando las coordenadas cartesianas con las q_i y $3N$ ecuaciones (19). El número de coordenadas independientes o el número de grados de libertad, como dicen los físicos, depende del sistema tratado y las restricciones en el movimiento.

principio supone el conocimiento de las energías cinética y potencial de un problema, no requiere el conocimiento de las fuerzas que actúan. Con su principio, Lagrange dedujo leyes básicas para la mecánica y solucionó muchos problemas nuevos, aunque no era lo suficientemente amplio como para incluir todos los problemas de que se ocupa la dinámica. Su trabajo sobre el principio de acción está completamente discutido en su *Mécanique analytique*; también inició el movimiento para deducir las leyes de otras ramas de la física de principios variacionales análogos al de acción mínima. El mismo proporcionó un principio variacional para una clase más general de problemas hidrodinámicos. Este tema lo retomaremos en nuestro estudio de la aportación del siglo XIX.

Desde el punto de vista matemático, el trabajo de Lagrange sobre la acción mínima dio la mayor relevancia al cálculo de variaciones. En particular, Lagrange había derivado las ecuaciones de Euler para una integral cuyo integrando contiene una variable independiente pero varias variables dependientes y sus derivadas. Se trata de una extensión del cálculo original del problema de variaciones, el cual contiene únicamente una variable dependiente y su derivada. En este caso más general, las ecuaciones de Euler son un sistema de ecuaciones diferenciales de segundo orden en las q_i.

6. La segunda variación

La ecuación diferencial de Euler, como pensaron éste y Lagrange, es únicamente una condición necesaria de la solución para dar un máximo o un mínimo. Ellos aplicaron la ecuación diferencial a fin de encontrar la solución y entonces especularon sobre bases intuitivas o físicas si daba un máximo o un mínimo. El papel de la ecuación de Euler es del todo análogo a la condición $f'(x) = 0$ en el cálculo ordinario. Un valor de x que maximiza o minimiza $y = f(x)$ debe satisfacer $f'(x) = 0$, pero el recíproco no necesariamente es verdadero.

La pregunta de qué condiciones adicionales debe satisfacer una solución de la ecuación de Euler para suministrar de hecho un valor máximo o mínimo de una integral, dependiendo de la $y(x)$, fue atacado sin éxito en 1782 por Laplace y retomado por Legendre en 1786 [27]. Guiado por el hecho que en el cálculo ordinario el signo de

[27] *Hist. de l'Acad. des Sci., Paris*, 1786, 7-37, pub. 1788.

$f''(x)$ en un valor x para el cual $f'(x)=0$ determina si $f(x)$ tiene un máximo o un mínimo, Legendre consideró la segunda variación $\delta^2 J$, remodelada su forma, y concluyó que J es un máximo para la curva $y(x)$ que satisface la ecuación de Euler y pasa por (x_0, y_0) y (x_1, y_1) con tal que $f_{y'y'} \leq 0$ en cada x a lo largo de $y(x)$. Del mismo modo, J es un mínimo para una $y(x)$ satisfaciendo las primeras dos condiciones siempre que $f_{y'y'} \geqslant 0$ en cada x a lo largo de $y(x)$. Más adelante, Legendre extendió este resultado a integrales más generales que (6). Sin embargo, Legendre se dio cuenta en 1787 que la condición sobre $f_{y'y'}$ era únicamente una condición necesaria sobre $y(x)$ para que sea una curva maximizante o minimizante. El problema de encontrar condiciones suficientes de que una curva $y(x)$ maximice o minimice una integral tal como (6) no fue resuelto en el siglo XVIII.

Bibliografía

Bernoulli, Jacques: *Opera*, 2 vols. 1744, reimpreso por Birkhauser, 1968.

Bernoulli, Jean: *Opera Omnia*, 4 vols., 1742, Georg Olms (reimpreso), 1968.

Bliss, Gilbert A.: *The Calculus of Variations*, Open Court, 1925.

—: «The Evolution of Problems in the Calculus of Variations», *Amer. Math. Monthly*, 43, 1936, 598-609.

Cantor, Moritz: *Vorlesungen über Geschichte der Mathematik*, B. G. Teubner, 1898 y 1924, vol. 3, cap. 117, y vol. 4, 1066-1074.

Caratheodory, C.: Introducción a las Series (1), vol. 24 de la *Opera Omnia* de Euler, viii-lxii, Orell Fussli, 1952. También en C. Caratheodory: *Gesammelte mathematische Schriften*, C. H. Beck, 1957, vol. 5, pp. 107-174.

Darboux, Gaston: *Leçons sur la théorie générale des surfaces*, 2.ª ed., Gauthier-Villars, 1914, vol. 1, libro III, caps. 1-2.

Euler, Leonhard: *Opera Omnia*, (1), vols. 24-25, Orell Fussli, 1952.

Hofmann, Joseph E.: «Über Jakob Bernoullis Beiträge zur Infinitesimalmathematik», *L'Enseignement mathématique*, (2), 2, 1956, 61-171; publicado por separado por el Institut de Mathématiques, Ginebra, 1957.

Huke, Aline: *An Historical and Critical Study of the Fundamental Lemma in the Calculus of Variations*, University of Chicago Contributions to the Calculus of Variations, Chicago University Press, vol. 1, 1930, pp. 45-160.

Lagrange, Joseph Louis: *Œuvres de Lagrange*, Gauthiers-Villars, 1867-69, ensayos relevantes en los volúmenes 1 a 3.

—: *Mécanique analytique*, 2 vols, 4.ª ed. Gauthier-Villars, 1889.
Lecat, Maurice: *Bibliographie du calcul des variations depuis les origines jusqu'à 1850*, Gante, 1916.
Montucla, J. F.: *Histoire des mathématiques*, 1802, Albert Blanchard (reimpresión), 1960, vol. 3, 643-658.
Porter, Thomas Isaac: «A History of the Classical Isoperimetric Problem», *University of Chicago Contributions to the Calculus of Variations*, Chicago University Press, 1933, vol. 2, 475-517.
Smith, David E.: *A Source Book in Mathematics*, Dover (reimpresión), 1959, pp. 644-655.
Struik, D. J.: *A Source Book in Mathematics (1200-1800)*, Harvard Unviersity Press, 1969, pp. 391-413.
Todhunter, Isaac: *A History of the Calculus of Variations during the Nineteenth Century*, 1861, Chelsea (reimpresión), 1962.
Woodhouse, Robert: *A History of the Calculus of Variations in the Eighteenth Century*, 1810, Chelsea (reimpresión), 1964.

Capítulo 25
EL ALGEBRA DEL SIGLO XVIII

Yo presento el análisis superior como se encontraba en su niñez, pero usted lo está llevando a su madurez.

JEAN BERNOULLI, en una carta a Euler

1. La situación de los números

El significado del concepto de límite en el siglo XVIII era aún oscuro y difícilmente se distinguía entre álgebra y análisis, por lo que, en nuestro actual enfoque, sería conveniente separar estos dos campos de actividades. En el siglo XVII, el álgebra fue un centro importante de interés; en el XVIII se convirtió en algo subordinado al análisis; exceptuando la teoría de números, la motivación para trabajar en ella vino en gran parte del análisis.

Ya que la base del álgebra es el sistema numérico, describamos el estado de esta materia. En 1700 todos los miembros familiares del sistema —números enteros, fraccionarios, irracionales, y números negativos y complejos— eran conocidos. Sin embargo, la oposición hacia los tipos más nuevos se manifestó a lo largo del siglo. Típicas son las objeciones del matemático inglés barón Francis Maseres (1731-1824), miembro del Clare College en Cambridge y miembro de la Royal Society. Maseres, quien escribió ensayos aceptables en matemáticas y un tratado sustancial sobre la teoría de seguros de vida, publicó en 1759 su *Dissertation on the Use of the Negative Sign in Algebra (Disertación sobre el uso del signo negativo en álge-*

bra), donde muestra cómo evitar los números negativos (excepto para indicar la sustracción de una cantidad mayor de una menor), y en particular las raíces negativas, mediante la cuidadosa segregación de tipos de ecuaciones cuadráticas de tal forma que aquellas con raíces negativas son consideradas separadamente; y, por supuesto, las raíces negativas deben ser rechazadas. El hace lo mismo con las cúbicas. Más adelante dice de las raíces negativas:

[...] sirven únicamente, en lo que yo puedo juzgar, para confundir toda la doctrina de ecuaciones y para volver en cosas oscuras y misteriosas las que son en su propia naturaleza excesivamente simples y ordinarias [...]. Se debería desear, por lo tanto, que las raíces negativas nunca hubieran sido admitidas dentro del álgebra o que fueran, de nuevo, descartadas de ella; ya que, si esto fuera hecho, hay razón de sobra para imaginar que las objeciones que muchos hombres cultos e ingeniosos ahora hacen de los cómputos algebraicos, como ser oscurecidos y confundidos con nociones casi ininteligibles, serían por consiguiente suprimidas; siendo inevitable que el álgebra o la aritmética universal es, por su propia naturaleza, una ciencia no menos simple, clara y capaz de demostración que la geometría.

Ciertamente, los números negativos no fueron realmente bien comprendidos hasta los tiempos modernos, Euler, en la última mitad del siglo XVIII, aún creía que los números negativos eran mayores que el ∞. También discutió que $(-1)\cdot(-1)=-1$ porque el producto debe ser $+1$ o -1 y ya que $1\cdot(-1)=-1$, entonces $(-1)\cdot(-1)=+1$. Carnot, el famoso geómetra francés, pensó que el uso de números negativos llevaba a conclusiones erróneas. Tan tarde como 1831, Augustus De Morgan (1805-710), profesor de matemáticas en el University College de Londres, famoso matemático y con trabajos en álgebra, en su *On the Study and Difficulties of Mathematics (Sobre el estudio y dificultades de las matemáticas)* dijo: «La expresión imaginaria $\sqrt{-a}$ y la expresión negativa $-b$ tienen este parecido: que cualquiera de ellas, cuando aparece como solución de un problema, indica alguna inconsistencia o absurdo. En cuanto se refiere al significado real, ambas son igualmente imaginarias, ya que $0-a$ es tan inconcebible como $\sqrt{-a}$.»

De Morgan ilustró esto por medio de un problema. Un padre tiene cincuenta y seis años; su hijo, veintinueve. ¿Cuándo será el padre el doble de viejo que su hijo? El resuelve $56+x=2(29+x)$ y obtiene $x=-2$. Así el resultado, dice, es absurdo. Pero, continúa, si cambiamos x por $-x$ y solucionamos $56-x=2(29-x)$, obtene-

mos $x = 2$. Concluye que expresamos el problema original equivocadamente y fuimos, por tanto, conducidos a una respuesta inaceptable negativa. De Morgan insistió en que era absurdo considerar número menores que cero.

A pesar de que en el siglo XVIII no se hizo nada por aclarar el concepto de número irracional, hubo cierto progreso en esta materia. En 1737, Euler mostró, sustancialmente, que e y e^2 eran irracionales y Lambert mostró que π es irracional (cap. 20, sec. 6). El trabajo sobre la irracionalidad de π estuvo motivado en gran parte por el deseo de resolver el problema de la cuadratura del círculo. La conjetura de Legendre de que π podría no ser la raíz de una ecuación algebraica con coeficientes racionales, llevó a la distinción entre tipos de irracionales. Cualquier raíz, real o compleja, de cualquier ecuación algebraica (polinomial) con coeficientes racionales es llamada número algebraico. Así que a las raíces de

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}\,x + a_n = 0,$$

donde los a_i son números racionales se les llama números algebraicos. Consecuentemente, todo número racional y algunos irracionales son números algebraicos, ya que cualquier número racional c es la raíz de $x - c = 0$ y $\sqrt{2}$ es una raíz de $x^2 - 2 = 0$. Aquellos números que no son algebraicos se denominan trascendentes porque, como indicó Euler, «ellos trascienden el poderío de los métodos algebraicos». Esta distinción entre números algebraicos y trascendentales, al menos ya era reconocida por Euler en fechas tan tempranas como 1744. Euler conjeturó que el logaritmo de una base racional debía ser racional o trascendente. Sin embargo, ningún número trascendente era conocido en el siglo XVIII y el problema de mostrar que había números trascendentes permanecía abierto.

Los números complejos eran más que un veneno para los matemáticos del siglo XVIII. De Cardano hasta cerca de 1700, prácticamente se ignoró estos números. Entonces (cap. 19, sec. 3), los números complejos se utilizaron para hallar integrales por el método de descomposición en fracciones, al cual siguió una extensa controversia acerca de números complejos y los logaritmos de números negativos y complejos. A pesar de la correcta resolución del problema de los logaritmos de los números complejos, ni Euler ni otros matemáticos tenían ideas claras sobre tales números.

Euler intentó comprender lo que son realmente los números com-

plejos, y en su *Vollstandige Anleitung zur Algebra (Introducción completa al álgebra)*, que primero apareció en ruso en 1768-69 y en alemán en 1770, y que es el mejor texto de álgebra del siglo XVIII, dice:

[...]. Porque todos los números concebibles, o son mayores que cero o menores que cero o iguales a cero, entonces es claro que las raíces cuadradas de números negativos no pueden estar incluidas entre los posibles números [números reales]. Consecuentemente, debemos decir que éstos son números imposibles. Y dicha circunstancia nos lleva al concepto de tales números, los cuales por su naturaleza son imposibles, y ordinariamente son llamados números imaginarios o fantasiosos, ya que sólo existen en la imaginación.

Euler cometió errores con los números complejos. En su *Algebra* escribe $\sqrt{-1} \cdot \sqrt{-4} = \sqrt{4} = 2$, ya que $\sqrt{a}\,\sqrt{b} = \sqrt{ab}$. También da $i^i = 0{,}2078795763$, pero olvida los otros valores de esta cantidad. Este número lo proporcionó originalmente en una carta a Goldbach de 1746 y también en su ensayo de 1749 sobre la controversia entre Leibniz y Jean Bernoulli (cap. 19, sec. 3). A pesar de que llama a los números complejos números imposibles, Euler dice que pueden usarse cuando atacamos problemas acerca de los cuales ignoramos si tienen o no una respuesta. De este modo, si preguntamos cómo separar 12 en dos partes cuyo producto sea 40, deberíamos encontrar que las partes son $6 + \sqrt{-4}$ y $6 - \sqrt{-4}$. Por lo que, dice, reconocemos que este problema no puede ser resuelto.

Fueron realizados algunos pasos positivos más allá de la correcta conclusión de Euler sobre los logaritmos de números complejos, pero su influencia en el siglo XVIII fue limitada. En su *Algebra* (1685, caps. 66-69), John Wallis mostró cómo representar geométricamente las raíces complejas de una ecuación cuadrática con coeficientes reales. Wallis dijo, en efecto, que los números complejos no son más absurdos que los números negativos; y si estos últimos pueden ser representados en una línea recta, entonces es posible representar los números complejos en un plano. Empezó con un eje sobre el cual los números reales fueron marcados en relación con un origen; una distancia a partir de este origen a lo largo del eje representaba la parte real de la raíz, siendo medida la distancia en la dirección positiva o negativa del eje según si dicho número era positivo o negativo. Desde el punto sobre el eje de los reales así determinado, era trazada una línea perpendicular al eje de los reales cuya longitud

representaba el número que, multiplicado por $\sqrt{-1}$, proporcionaba la parte imaginaria de la raíz, siendo la línea trazada en una dirección o la opuesta dependiendo si el número era positivo o negativo. (Fracasó al introducir el mismo eje y como el eje de los imaginarios.) Wallis procedió entonces a dar una construcción geométrica de las raíces de $ax^2 + bx + c = 0$ cuando las raíces son reales y cuando son complejas. Su trabajo fue correcto, pero no fue una representación útil de $x + iy$ con otros propósitos. Hubo otros intentos para representar los números complejos geométricamente en el siglo XVIII, pero no fueron demasiado útiles. Tampoco las representaciones geométricas de este tiempo hicieron a los números complejos más aceptables.

En la parte inicial del siglo, los matemáticos pensaban que las raíces diferentes de los números complejos introducirían tipos diferentes (u órdenes) de números complejos y que podía haber raíces ideales cuya naturaleza no podían especificar pero que de alguna manera podían ser calculadas. Pero D'Alembert, en su trabajo premiado *Réflexions sur la cause générale des vents* (*Reflexiones sobre la causa general de los vientos*, 1747), afirmó que cada expresión construida sobre los números complejos por medio de operaciones algebraicas (en las cuales incluyó elevarlos a una potencia arbitraria) es un número complejo de la forma $A + B\sqrt{-1}$. La gran dificultad que encontró para probar esta aserción fue el caso de $(a + bi)^{g+hi}$. Su demostración de este hecho tuvo que ser enmendada por Euler, Lagrange y otros. En su *Encyclopédie*, D'Alembert mantuvo un silencio discreto acerca de los números complejos.

A lo largo de todo el siglo XVIII, los números complejos fueron usados de una manera lo suficientemente efectiva por los matemáticos como para adquirir cierta confianza en ellos (cap. 19, sec. 3; cap. 27, sec. 2). Cuando, usados en estados intermedios de los argumentos matemáticos, los resultados fueron correctos, este hecho tuvo un efecto notable. Aun así, había dudas en cuanto a la validez de los argumentos y frecuentemente también de los resultados.

En 1799, Gauss dio la primera prueba del teorema fundamental del álgebra y, como dependía esencialmente del reconocimiento de los números complejos, éste fortaleció la posición de estos números. El siglo XIX se precipitó entonces hacia adelante ciegamente con las funciones complejas. Pero aún mucho más tarde de que esta teoría fuese desarrollada y empleada en hidrodinámica, los profesores de Cambridge conservaron «una repulsión invisible a la objetable $\sqrt{-1}$,

y complicados utensilios fueron adoptados para evitar su aparición o uso siempre que fuera posible».

La actitud general hacia los números complejos, incluso hasta 1831, puede verse en el libro de De Morgan *On the Study and Difficulties of Mathematics (Sobre el estudio y las dificultades de las matemáticas)*. De Morgan comenta que este libro no contiene nada que no pueda ser encontrado en los mejores trabajos de entonces en Oxford y Cambridge. Volviendo a los números complejos, dice:

> [...] hemos mostrado que el símbolo $\sqrt{-a}$ carece de significado, o mejor es autocontradictorio y absurdo. Sin embargo, por medio de tales símbolos, se establece una parte del álgebra que es de gran utilidad. Ello depende del hecho, que debe ser verificado por la experiencia, de que las reglas comunes del álgebra pueden ser aplicadas a estas expresiones [números complejos] sin llevarnos a resultados falsos. Una llamada a la experiencia de esta naturaleza parece ser contraria a los primeros principios establecidos al comienzo del presente trabajo. Nosotros no podemos negar que así sea en la realidad, pero debe ser recordado que ésta no es más que una parte pequeña y aislada de una materia inmensa, en todas las demás ramas de la cual se aplican sus principios en toda su extensión.

Los «principios» a los que se refiere son que las verdades matemáticas deben ser deducidas a partir de axiomas por medio del método deductivo.

Más adelante compara las raíces negativas con las raíces complejas.

> Existe, entonces, esta diferencia distintiva entre los resultados negativos y los imaginarios. Cuando la respuesta a un problema es negativa, cambiando el signo de x en la ecuación que produjo el resultado, podemos o descubrir un error en el método de formación de la ecuación o mostrar que la cuestión del problema es demasiado limitada, y que puede ser extendida hasta admitir una respuesta satisfactoria. Cuando la respuesta a un problema es imaginaria éste no es el caso [...]. Nosotros debemos abocarnos a detener el progreso del estudiante entrando en todos los argumentos en pro y en contra de tales cuestiones, como el uso de cantidades negativas, etc., las cuales no podía él entender, y que son inconclusas en ambos sentidos; pero se le puede advertir que existe una dificultad, la naturaleza de la cual se le debe señalar, y entonces tal vez pueda, tomando en cuenta un número suficiente de ejemplos, tratados por separado, adquirir confianza en los resultados a los cuales llevan las reglas.

Para cuando De Morgan escribió estas líneas, los conceptos de nú-

mero complejo y función compleja estaban en el camino correcto para su clarificación. Pero la difusión del nuevo conocimiento era lenta. Ciertamente, a lo largo del siglo XVIII y la primera parte del XIX, el significado de los números complejos fue discutido con énfasis. Todos los argumentos de Jean Bernoulli, D'Alembert y Euler estuvieron continuamente sujetos a discusión. Aun los libros de texto del siglo XX sobre trigonometría proporcionaron presentaciones que empleaban los números complejos con pruebas que evitaban $\sqrt{-1}$.

También debemos notar aquí otra cuestión cuya significación está casi en proporción inversa a la concisión con la que puede ser establecida: en el siglo XVIII nadie se preocupó por la lógica de los sistemas de números reales o complejos. Lo que Euclides había hecho en el libro quinto de los *Elementos* para establecer las propiedades de las magnitudes inconmesurables fue evitado. El que esta exposición hubiera estado ligada a la geometría, mientras que ahora la aritmética y el álgebra eran independientes de la geometría, explica en parte su abandono. Más aún, este desarrollo lógico, aunque fuera modificado para liberarlo de la geometría, no podía establecer los fundamentos lógicos de los números negativos y complejos; y, también, este hecho pudo haber causado que los matemáticos desistieran de cualquier intento por fundamentar rigurosamente el sistema numérico. Finalmente, el siglo dedicaba su interés al uso de las matemáticas en las ciencias, y como las reglas de operación eran intuitivamente seguras, al menos para los números reales, nadie se preocupó realmente por los fundamentos. Típico es el argumento de D'Alembert en su ensayo sobre números negativos en la *Encyclopédie*. EL ensayo no es del todo claro y D'Alembert concluye que «las reglas algebraicas de operación con números negativos son admitidas generalmente por todos y reconocidas como exactas, cualquiera que sea la idea que tengamos sobre estas cantidades». Los diversos tipos de números, nunca introducidos adecuadamente en el mundo, ganaron sin embargo un lugar firme en la comunidad matemática de ese siglo.

2. La teoría de ecuaciones

Una de las investigaciones que continuó a partir del siglo XVII, acaso con una ligera interrupción, fue la solución de ecuaciones po-

linómicas. El tema es fundamental en matemáticas y de igual manera el interés por obtener mejores métodos para resolver ecuaciones de cualquier grado, obteniendo mejores métodos para aproximar las raíces de ecuaciones, y el completar la teoría resultaba natural —en particular, demostrando que cualquier ecuación polinómica de grado n-ésimo tiene n raíces. Además, el uso del método de descomposición en fracciones para la integración originó la cuestión de si cualquier polinomio con coeficientes reales podía ser descompuesto en un producto de factores lineales o un producto de un factor lineal y otro cuadrático con coeficientes reales, para evitar el uso de números complejos.

Como vimos en el capítulo 19 (sec. 4), Leibniz no creía que cualquier polinomio con coeficientes reales podía ser descompuesto en factores lineales y cuadráticos con coeficientes reales. Euler tomó la postura correcta. En una carta a Nicolas Bernoulli (1687-1759) del 10 de octubre de 1742, Euler afirmó, sin dar una demostración, que un polinomio de grado arbitrario con coeficientes reales se podía expresar de esa manera. Nicolas no creyó que la aserción fuera correcta y dio un ejemplo

$$x^4 - 4x^3 + 2x^2 + 4x + 4$$

con los ceros $1 + \sqrt{2 + \sqrt{-3}}$, $1 - \sqrt{2 + \sqrt{-3}}$, $1 - \sqrt{2 - \sqrt{-3}}$, $1 + \sqrt{2 - \sqrt{-3}}$, los cuales según él contradicen la afirmación de Euler. El 15 de diciembre de 1742, escribiendo a Goldbach (Fuss, vol. 1, 169-171), señaló que las raíces complejas aparecen en pares conjugados, de tal manera que el producto $x - (a + b\sqrt{-1})$ y $x - (a - b\sqrt{-1})$, donde $a + b\sqrt{-1}$ y $a - b\sqrt{-1}$ son conjugados, proporciona una expresión cuadrática con coeficientes reales. Euler, entonces, mostró que esto era cierto para el ejemplo de Bernoulli. Pero también Goldbach rechazó la idea de que cada polinomio con coeficientes reales podía ser factorizado en factores reales y dio el ejemplo $x^4 + 72x - 20$. Euler entonces le mostró a Goldbach que este último había cometido un error y que él (Euler) había probado su teorema para polinomios hasta de grado seis. Sin embargo, Goldbach no se convenció, ya que Euler no había tenido éxito en dar una prueba general de su aserción.

El meollo del problema de factorizar un polinomio real en factores lineales y cuadráticos con coeficientes reales era probar que cada polinomio tenía al menos una raíz real o compleja. La prueba

de este hecho, llamado teorema fundamental del álgebra, se convirtió en una finalidad primordial.

Las pruebas ofrecidas por D'Alembert y Euler fueron incompletas. En 1772 [1] Lagrange, con un argumento largo y detallado, «completó» la prueba de Euler. Pero Lagrange, al igual que Euler y sus contemporáneos, aplicaron libremente las propiedades ordinarias de los números a lo que supuestamente eran las raíces, sin establecer que las raíces, en el peor de los casos, debían ser números complejos. Debido a que la naturaleza de las raíces era desconocida, la prueba, de hecho, también era incompleta.

La primera demostración sustancial del teorema fundamental, pero no aceptable para los niveles modernos de rigor, fue dada por Gauss en su tesis doctoral de 1799, en Helmstadt [2]. En este texto critica los trabajos de D'Alembert, Euler y Lagrange y ofrece su propia prueba. El método de Gauss no fue calcular una raíz, sino demostrar su existencia. Gauss señalaba que las raíces complejas $a + ib$ de $P(x + iy) = 0$ corresponden a puntos (a, b) del plano y si $P(x + iy) = u(x, y) + iv(x, y)$, entonces (a, b) debe ser la intersección de las curvas $u = 0$ y $v = 0$. Mediante un estudio cualitativo de las curvas, muestra que un arco continuo de una une puntos de dos regiones distintas separadas por la otra. Entonces, la curva $u = 0$ debe cortar a la curva $v = 0$. El argumento era altamente original. Sin embargo, dependía de las gráficas de estas curvas, que eran algo complicadas, para mostrar que debían cruzarse. En este mismo ensayo, mostró que un polinomio de grado n-ésimo podía ser expresado como el producto de factores lineales y cuadráticos con coeficientes reales.

Gauss dio tres pruebas más del teorema. En la segunda prueba [3] evitó el uso de argumentos geométricos. Aquí también demostró que el producto de las diferencias de cada dos raíces (las que nosotros, siguiendo a Sylvester, llamamos el discriminante) puede ser expresado como una combinación lineal del polinomio y de su derivada, de tal manera que una condición necesaria y suficiente para que el polinomio y su derivada tengan una raíz común es que el discriminante se anule. Sin embargo, esta segunda prueba suponía que un polinomio no puede cambiar de signo entre dos valores diferentes

---

[1] *Nouv. Mém. de l'Acad. de Berlin*, 1772, 222 y sigs. = *Œuvres*, 3, 479-516.

[2] *Werke*, 3, 1-30; también reproducido en Euler, *Opera*, (1), 6, 151-169.

[3] *Comm. Soc. Goth*., 3, 1814/15, 107-142 = *Werke*, 3, 33-56.

de x sin anularse entre ellos. La prueba de este hecho iba más allá del rigor de aquella época.

La tercera prueba [4] aplica de hecho lo que llamamos el teorema de la integral de Cauchy (cap. 27, sec. 4) [5]. La cuarta prueba [6] es una variación de la primera en cuanto se refiere al método. Sin embargo, en esta demostración Gauss usa números complejos con mayor libertad porque —como dice— ahora son de conocimiento común. Vale la pena notar que el teorema no se demostraba en toda su generalidad en muchas pruebas. Las primeras tres pruebas de Gauss y demostraciones posteriores de Cauchy, Jacobi y Abel supusieron que los coeficientes (literales) representaban números reales, mientras que el teorema en su totalidad incluye el caso de los coeficientes complejos. La cuarta demostración de Gauss sí permitía que los coeficientes del polinomio fueran números complejos.

El método de Gauss para el teorema fundamental del álgebra inauguró un nuevo acercamiento a la cuestión de la existencia matemática. Los griegos, sabiamente, habían reconocido que la existencia de las entidades matemáticas debía ser establecida aún antes que los teoremas acerca de ellas pudieran ser obtenidos. Su criterio de existencia era la constructibilidad. En el trabajo formal más explícito de los siglos posteriores, la existencia se establecía obteniendo o estableciendo la cantidad en cuestión. Por ejemplo, la existencia de solución de la ecuación cuadrática queda establecida al mostrar cantidades que satisfacen la ecuación. Pero en el caso de ecuaciones de grado superior a cuatro, este método no puede usarse. Por supuesto, una prueba de existencia tal como la de Gauss puede no ser una ayuda al calcular el objeto cuya existencia está siendo establecida.

Mientras se abría camino al trabajo que finalmente mostró cómo toda ecuación polinómica con coeficientes reales tiene una raíz, los matemáticos también luchaban para resolver ecuaciones de grado superior a cuatro usando procesos algebraicos. Leibniz y su amigo Tschirnhausen fueron los primeros en hacer serios esfuerzos. Leib-

---

[4] *Comm. Soc. Gott.*, 3, 1816 = *Werke*, 3, 59-64.

[5] Para una discusión de la tercera demostración de Gauss véase M. Bocher, «Gauss's Third Proof of the Fundamental Theorem of Algebra», *Amer. Math. Soc. Bull.*, 1, 1895, 205-209. Una traducción al inglés de la tercera prueba se encuentra en H. Meschkowski, *Ways of thought of great mathematicians*, Holden-Day, 1964.

[6] *Abhand. der Ges. der Wiss. zu Gött.*, 4, 1848/50, 3-34 = *Werke*, 3, 73-102.

niz [7] reconsideró el caso irreducible de la ecuación de tercer grado y se convenció de que no podía evitarse el uso de números complejos para resolver este caso. Entonces atacó la solución de la ecuación de quinto grado, pero sin éxito. Tschirnhausen [8] pensó que él había resuelto este problema al transformar la ecuación dada en una nueva por medio de una transformación $y = P(x)$, donde $P(x)$ es un polinomio adecuado de cuarto grado.

Esta transformación eliminó todas las potencias menos la x^5 y los términos constantes de la ecuación. Pero Leibniz mostró que para determinar los coeficientes de $P(x)$ había que resolver ecuaciones de grado superior a cinco, por lo que el método resultaba inútil.

Por un tiempo, el problema de hallar la solución de la ecuación de grado n-ésimo se centró en el caso especial $x^n - 1 = 0$, llamado ecuación binomial. Cotes y De Moivre mostraron, usando números complejos, que la solución de este problema se reduce a la división de la circunferencia de un círculo en n partes iguales. Para obtener las raíces por radicales (las soluciones trigonométricas no son necesariamente algebraicas) es suficiente considerar el caso de n primo impar, ya que si $n = pm$, donde p es primo, entonces se puede considerar $(x^m)^p - 1 = 0$. Si esta ecuación puede ser resuelta para x^m, entonces se considera $x^m - A$, donde A es cualquiera de las raíces de la ecuación precedente. Alexandre Théophile Vandermonde (1735-96) afirmó en su ensayo de 1771 [9] que toda ecuación de la forma $x^n - 1 = 0$, donde n es primo, es soluble por radicales. Sin embargo, Vandermonde únicamente verificó que esto es así para valores de n hasta 11. El trabajo decisivo en ecuaciones binominales fue hecho por Gauss (cap. 31, sec. 2).

El mayor esfuerzo hacia la solución de ecuaciones de grado superior a 4 se concentró en la ecuación general y, con este fin, cierto trabajo subsidiario sobre funciones simétricas resultó ser importante. La expresión $x_1x_2 + x_2x_3 + x_3 x_1$ es una función simétrica de x_1, x_2 y x_3, ya que el reemplazo de cualquier x_i por una x_j y x_j por x_i deja la expresión inalterada por completo. El interés por funciones simétricas surgió cuando los algebristas del siglo XVII notaron que Newton probó que las diversas sumas de los productos de las raíces de

[7] *Der Briefwechsel von Gottfried Wilhelm Leibniz mit Mathematikern*, Georg Olms (reimpresión), 1961, vol. 1, 547-564.

[8] *Acta Erud.*, 2, 1683, 204-207.

[9] *Hist. de l'Acad. des Sci., Paris*, 1771, 365-416, pub. 1774.

una ecuación polinomial pueden ser expresados en términos de los coeficientes. Por ejemplo, para $n = 3$, la suma de los productos tomados dos a dos,

$$a_1a_2 + a_2a_3 + a_3a_1,$$

es una función simétrica elemental; y si la ecuación es escrita como

$$x^3 - c_1x^2 + c_2x - c_3 = 0,$$

la suma es igual a c_2. El progreso hecho en el ensayo de 1771 de Vandermonde fue mostrar que cualquier función simétrica de las raíces es posible expresarla en términos de los coeficientes de la ecuación.

El trabajo sobresaliente del siglo XVIII en el problema de la solución de ecuaciones por radicales, después de los grandes esfuerzos de hombres como Euler [10], fue el de Vandermonde —en su ensayo de 1771— y el de Lagrange, en su amplio artículo «Reflexions sur la résolution algébrique des équations» («Reflexiones sobre la resolución algebraica de ecuaciones») [11]. Las ideas de Vandermonde son similares pero no tan extensas ni tan claras. Nosotros debemos, por tanto, presentar la versión de Lagrange. Este se propuso a sí mismo el objetivo de analizar los métodos de solución de las ecuaciones de tercero y cuarto grado, para ver por qué éstos resultaban bien y para ver qué pista le proporcionaban estos métodos en la solución de ecuaciones de grados superiores.

Para la ecuación de tercer grado

$$x^3 + nx + p = 0 \tag{1}$$

Lagrange notó que si se hace la transformación (cap. 13, sec. 4)

$$x = y - (n/3y), \tag{2}$$

se obtiene la ecuación auxiliar

$$y^6 + py^3 - n^3/27 = 0. \tag{3}$$

---

[10] *Comm. Acad. Sci. Petrop.*, 6, 1732/3, 216-231, pub. 1738 = *Opera*, (1), 6, 1-19, y *Novi Comm. Acad. Sci. Petrop.*, 9, 1762/63, 70-98, pub. 1764 = *Opera*, (1), 6, 170-196.

[11] *Nouv. Mém. de l'Acad de Berlin*, 1770, 134-215, pub. 1772 y 1771, 138-254, pub. 1773 = *Œuvres*, 3, 205-421.

Esta ecuación también se conoce como la ecuación reducida, porque es cuadrática en y^3 y con $r = y^3$ se convierte en

$$r^2 + pr - n^3/27 = 0. \qquad (4)$$

Ahora vemos que las raíces r_1 y r_2 de esta ecuación pueden calcularse en términos de los coeficientes de la original y para volver a y partiendo de r, debemos introducir raíces cúbicas o resolver

$$y^3 - r = 0.$$

Entonces, si hacemos que ω sea la raíz cúbica de unidad $(-1 + \sqrt{-3})/2$ los valores de y son

$$\sqrt[3]{r_1},\ \omega\sqrt[3]{r_1},\ \omega^2\sqrt[3]{r_1},\ \sqrt[3]{r_2},\ \omega\sqrt[3]{r_2},\ \omega^2\sqrt[3]{r_2}$$

y las distintas soluciones de (1) son

$$x_1 = \sqrt[3]{r_1} + \sqrt[3]{r_2}, \qquad x_2 = \omega\sqrt[3]{r_1} + \omega^2\sqrt[3]{r_2}, \qquad x_3 = \omega^2\sqrt[3]{r_1} + \omega\sqrt[3]{r_2},$$

por lo que las soluciones de la ecuación original se obtienen en términos de las de la ecuación reducida.

Lagrange mostró que todos los diferentes procedimientos aplicados por sus predecesores equivalen al método anterior. Señaló que es necesario volver nuestra atención no a x como una función de los valores de y, sino a y como una función de x porque es la ecuación reducida la que permite la solución: el secreto debe estar en la relación que expresa las soluciones de la ecuación reducida en términos de las soluciones de la ecuación propuesta.

Lagrange notó que cada uno de los valores de y puede escribirse (ya que $1 + \omega + \omega^2 = 0$) en la forma

$$y = \frac{1}{3}(x_1 + \omega x_2 + \omega^2 x_3) \qquad (5)$$

cuando x_1, x_2 y x_3 son tomados en órdenes particulares. Un examen de esta expresión hace percibir dos propiedades de la ecuación reducida en y. Primero, las raíces x_1, x_2 y x_3 en la expresión para y no son una sola opción fija para x_1, x_2 y x_3, y la expresión es, por así decirlo, ambigua. De este modo cualquiera de los tres valores de x

puede ser x_1, cualquiera de los otros dos x_2, etc. Pero hay 3! permutaciones de las x_i. Por lo cual hay seis valores para y e y debe satisfacer una ecuación de sexto grado. Así, el grado de la ecuación reducida está determinado por el número de permutaciones de las raíces de la ecuación propuesta.

En segundo lugar, la relación (5) también muestra por qué es factible reducir la ecuación de sexto grado reducida a la ecuación de segundo grado, dado que entre las seis permutaciones, tres (incluyendo la identidad) vienen de intercambiar las x_i y tres de intercambiar únicamente dos manteniendo una fija. Pero, entonces, en función del valor ω, los seis valores de y que resulten están relacionados por

$$y_1 = \omega^2 y_2 = \omega y_3; \qquad y_4 = \omega^2 y_5 = \omega y_6 \tag{6}$$

y al elevar al cubo,

$$y_1^3 = y_2^3 = y_3^3, \qquad y_4^3 = y_5^3 = y_6^3.$$

Otra manera de expresar este resultado es decir que la función

$$(x_1 + \omega x_2 + \omega^2 x_3)^3$$

puede tomar únicamente dos valores bajo la permutación de x_1, x_2 y x_3, y ésta es la razón de que la ecuación que satisface y sea cuadrática en y^3. Más aún, los coeficientes de la ecuación de sexto grado que satisface y son funciones racionales de los coeficientes de la cúbica original.

En el caso de la ecuación general de cuarto grado en x, Lagrange considera

$$y = x_1 x_2 + x_3 x_4.$$

Esta función de las cuatro raíces toma solamente tres valores distintos para las veinticuatro permutaciones posibles de las cuatro raíces. De aquí que ha de haber una ecuación de tercer grado que satisface y y los coeficientes de esta ecuación deben ser funciones racionales de los de la ecuación original. Estos argumentos se aplican a la ecuación de cuarto grado.

A continuación, Lagrange considera la ecuación general de grado n-ésimo

$$x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0. \qquad (7)$$

Los coeficientes de esta ecuación son supuestamente independientes; no debe existir una relación entre los a_1, por tanto, las raíces también deben ser independientes, ya que si hubiera una relación entre las raíces sería posible mostrar que esto debería ser cierto respecto a los coeficientes (ya que, esencialmente, los coeficientes son funciones simétricas de las raíces). Así, las n raíces de la ecuación general han de considerarse como variables independientes, y cada función de ellas es una función de las variables independientes.

Para entender el plan de Lagrange para hallar la solución, consideremos primero

$$x^2 + bx + c = 0.$$

Conocemos dos funciones de las raíces, a saber: $x_1 + x_2$ y x_1x_2 que son funciones simétricas, lo cual significa que no cambian cuando los papeles de las raíces son cambiados. Cuando una función no cambia por la permutación llevada a cabo en sus variables, se dice que la función admite la permutación. Por ejemplo, la función $x_1 + x_2$ admite la permutación de x_1 y x_2, pero no así la función $x_1 - x_2$.

Lagrange demostró entonces dos proposiciones importantes: si una función $\phi(x_1, x_2, \ldots, x_n)$ de las raíces de la ecuación general de grado n admite todas las permutaciones de las x_i que admite otra función $\psi(x_1, x_2, \ldots, x_n)$ (y posiblemente otras permutaciones que ψ no admite), la función ϕ puede ser expresada racionalmente en términos de ψ y los coeficientes de la ecuación general (7). Así, la función x_1 de la ecuación cuadrática admite todas las permutaciones que admite $x_1 - x_2$ (sólo existe una, la identidad), entonces

$$x_1 = \frac{-b + (x_2 - x_1)}{2}.$$

La prueba de Lagrange de tal proposición muestra también cómo expresar ϕ como una función racional de ψ.

La segunda proposición de Lagrange afirma: si una función $\phi(x_1, x_2, \ldots, x_n)$ de las raíces de la ecuación general no permite todas las

permutaciones admitidas por una función $\psi(x_1, x_2, \ldots, x_n)$, pero toma para las permutaciones que admite ψ r valores diferentes, entonces ϕ es una raíz de una ecuación de grado r cuyos coeficientes son funciones racionales de ψ y de la ecuación general dada de grado n. Esta ecuación de grado r puede ser construida. Así, $x_1 - x_2$ no admite todas las permutaciones de $x_1 + x_2$, pero toma los dos valores de $x_1 - x_2$ y $x_2 - x_1$; por tanto, $x_1 - x_2$ es una raíz de una ecuación de segundo grado cuyos coeficientes son funciones racionales de $x_1 + x_2$ y de b y c. De hecho, $x_1 - x_2$ es una raíz de

$$t^2 - (b^2 - 4c) = 0,$$

porque $b^2 - 4ac = (x_1 - x_2)^2$. Con el valor de esta raíz, a saber $\sqrt{b^2 - 4c}$, podemos encontrar x_1 por medio de la ecuación precedente para x_1.

Del mismo modo, para las ecuaciones $x^3 + px + q = 0$ la expresión (la ϕ)

$$(x_1 + \omega x_2 + \omega^2 x_3)^3,$$

donde $\omega = (-1 + i\sqrt{3})/2$, toma dos valores con las seis posibles permutaciones de las raíces, mientras que $x_1 + x_2 + x_3$ (la ψ) admite las seis permutaciones. Si los dos valores son denotados por A y B, se puede mostrar que A y B son raíces de una ecuación de segundo grado cuyos coeficientes son racionales en p y q (ya que $x_1 + x_2 + x_3 = 0$). Si resolvemos la ecuación de segundo grado y las raíces son A y B, podemos entonces encontrar x_1, x_2 y x_3 a partir de

$$\begin{gathered} x_1 + x_2 + x_3 = 0 \\ x_1 + \omega x_2 + \omega^2 x_3 = \sqrt[3]{A} \\ x_1 + \omega^2 x_2 + \omega x_3 = \sqrt[3]{B}. \end{gathered}$$

Para la ecuación de cuarto grado, Lagrange empezó con la función

$$x_1 x_2 + x_3 x_4, \tag{8}$$

que toma tres valores diferentes para las veinticuatro permutaciones posibles de las raíces, mientras que $x_1 + x_2 + x_3 + x_4$ admite todas las veinticuatro. De aquí que (8) sea una raíz de una ecuación de tercer grado cuyos coeficientes son funciones racionales de los de la ecuación original. Y, de hecho, la ecuación auxiliar o reducida para la ecuación general de cuarto grado es de tercer grado.

Para la ecuación de grado n-ésimo con coeficientes generales, Lagrange tuvo la idea de empezar con una función simétrica f_0 de las raíces que admiten todas las $n!$ permutaciones de las raíces. Tal función, señala, pudiera ser $x_1 + x_2 + \cdots + x_n$. Entonces él escogería una función ϕ_1 que admite únicamente algunas de las sustituciones. Supóngase que ϕ_1 toma para las $n!$ permutaciones posibles, digamos, r diferentes valores; ϕ_1 será entonces una raíz de una ecuación de grado r cuyos coeficientes son funciones racionales de ϕ_0 y los coeficientes de la ecuación general dada. La ecuación de grado r puede ser construida. Más aún, si se toma ϕ_0 como una de las funciones simétricas que relaciona raíces con coeficientes, entonces los coeficientes de la ecuación de grado r son conocidos por completo en términos de los coeficientes de la ecuación general dada. Si la ecuación de grado r pudiera ser resuelta algebraicamente, entonces ϕ_1 sería conocida en términos de los coeficientes de la ecuación original. Entonces se escoge una función ϕ_2 admitiendo sólo algunas de las permutaciones que admite ϕ_1. Puede ϕ_2 tomar, digamos, s valores diferentes bajo las sustituciones que admite ϕ_1. Por ello, ϕ_2 será una de las raíces de una ecuación de grado s cuyos coeficientes son funciones racionales de ϕ_1 y los coeficientes de la ecuación general dada. Los coeficientes de esta ecuación de grado s-ésimo serán conocidos si la ecuación de grado r-ésimo que tiene ϕ_1 como una de sus raíces puede resolverse. Si es posible resolver algebraicamente la ecuación de grado s-ésimo, entonces ϕ_2 será conocida en términos de los coeficientes de la ecuación original.

Continuamos así hasta ϕ_3, ϕ_4, . . . llegando a la función final, que tomamos como x_1. Si en este caso, las ecuaciones de grado r, s, . . . pueden ser resultas algebraicamente, se conocerá x_1 en términos de los coeficientes de la ecuación general dada. Las otras raíces x_2, x_3, . . ., x_n surgen del mismo proceso. Las ecuaciones de grado r, s, . . . son llamadas ahora ecuaciones resolventes [12].

El método de Lagrange fue útil para las ecuaciones generales de segundo, tercer y cuarto grado; aunque haya también intentado resolver la ecuación de quinto grado de esta manera, lo encontró tan difícil que lo abandonó. Así como para la cúbica únicamente se tenía que resolver una ecuación de segundo grado, para la de quinto grado

[12] Lagrange usó la palabra «resolvente» para formas especiales de las funciones ϕ_i y no para las ecuaciones que las ϕ_i satisfacen. Así, para la ecuación cúbica, $x_1 + \omega x_2 + \omega^2 x_3$ es una de las form as de la resolvente de Lagrange.

había que resolver una de sexto. En vano buscó Lagrange hallar una función resolvente (en su sentido del término) que verificara una ecuación de grado menor que cinco. Sin embargo, su trabajo no proporcionó criterio alguno para escoger ϕ_i que satisficieran ecuaciones resolubles algebraicamente. También, su método se aplicaba únicamente a la ecuación general porque sus dos proposiciones básicas suponen que las raíces son independientes.

Lagrange llegó a concluir que la solución de la ecuación general de grado superior (para $n > 4$) mediante operaciones algebraicas resultaba imposible con toda probabilidad (para ecuaciones especiales de grado superior ofreció poco Lagrange). Decidió que o el problema está más allá de las capacidades humanas o la naturaleza de las expresiones de las raíces debía ser diferente a todo lo que hasta ese tiempo se conocía. Gauss también, en sus *Disquisitiones* de 1801, declaró que el problema no podía resolverse.

El método de Lagrange, a pesar de su falta de éxito, proporciona una visión de la razones para los éxitos cuando $n \leq 4$ y los fracasos cuando $n > 4$; esta perspicacia fue capitalizada por Abel y Galois (cap. 31). Más aún, la idea de Lagrange de que uno debe considerar el número de valores que un función racional toma cuando sus variables son permutadas llevó a la teoría de grupos de permutaciones o de sustituciones. De hecho, tenía el teorema de que el orden de un subgrupo (el número de elementos) debe ser un divisor del orden del subgrupo. En el trabajo de Lagrange, que precedió cualquier otro trabajo en teoría de grupos, dicho resultado adopta la forma de que el número r de valores que toma ϕ_1 es un divisor de $n!$

Influenciado por Lagrange, Paolo Ruffini (1765-1822) —matemático, doctor, político y ardiente discípulo de Lagrange— hizo varios intentós durante los años 1799 a 1813 para probar que la ecuación general de grado superior al cuarto no podía ser resuelta algebraicamente. En su *Teoria generale delle equazioni (Teoría general de las ecuaciones)* [13], Ruffini tuvo éxito en demostrar, por el mismo método de Lagrange, que no existe ninguna ecuación resolvente (en el sentido de Lagrange) que satisfaga una ecuación de grado menor de cinco. De hecho, demostró que ninguna función racional de n elementos toma tres o cuatro valores bajo las permutaciones de los n elementos cuando $n > 4$. *Grosso modo*, intentó probar en sus *Riflessioni intorno alla soluzione delle equazioni algebrai-*

[13] 1799 = *Opere Mat.*, 1, 1-324.

che generali (Reflexiones en torno a la solución de la ecuación general algebraica) [14], que la solución algebraica de la ecuación general de grado $n > 4$ era imposible. Este esfuerzo no fue conclusivo, aunque, en un principio, Ruffini pensó que era correcto. Ruffini usó, pero no probó, el teorema auxiliar, conocido hoy como el teorema de Abel, de que si una ecuación es resoluble usando radicales, las expresiones para las raíces se dan en forma tal que los radicales en ellos sean funciones racionales con coeficientes racionales de las raíces de la ecuación dada y las raíces de unidad.

3. Determinantes y teoría de la eliminación

El estudio de un sistema de ecuaciones lineales, que nosotros escribimos ahora como

$$x_i = \sum_{n=1}^{n} a_{ij} y_j, \qquad i = 1, 2, \ldots, m \tag{9}$$

donde las x_i son conocidas y las y_j desconocidas, fue iniciado antes de 1678 por Leibniz. En 1693 [15], Leibniz usó un conjunto sistemático de índices para los coeficientes de un sistema de tres ecuaciones lineales en dos incógnitas x e y. Eliminó las dos incógnitas del sistema de las tres ecuaciones lineales y obtuvo un determinante, que actualmente denominamos el resultante del sistema. La anulación de este determinante expresa el hecho que hay una x y una y que satisfacen las tres ecuaciones.

La solución de ecuaciones lineales simultáneas de dos, tres y cuatro incógnitas por el método de determinantes fue creada por Maclaurin, probablemente en 1729, y publicada póstumamente en su *Treatise of Algebra* (*Tratado de Algebra*, 1748). Aunque no sea muy buena la notación, su regla es la que usamos hoy en día y que Cramer publicó en su *Introduction à l'analyse des lignes courbes algébriques* (*Introducción al análisis de líneas curvas algebraicas*, 1750). Cramer dio la regla para determinar los coeficientes de la cónica general, $A + By + Cx + Dy^2 + Exy + x^2 = 0$ pasando por cin-

[14] 1813 = *Opere Mat.*, 2, 155-268.
[15] *Math. Schriften*, 2, 229, 238-240, 245.

co puntos dados. Sus determinantes fueron, como en el presente, la suma de los productos formados al tomar uno y sólo un elemento de cada fila y columna, con el signo de cada producto determinado por el número de cambios de los elementos a partir de un orden fijado, siendo positivo el signo si este número era par o negativo si era impar. En 1764, Bezout [16] sistematizó el proceso de determinar los signos de los términos de un determinante. Dadas n ecuaciones lineales homogéneas con n incógnitas, Bezout demostró que la anulación del determinante de los coeficientes (la anulación del resultante) es una condición para que existan soluciones diferentes de cero.

Vandermonde [17] fue el primero en dar una exposición coherente y lógica de la teoría de los determinantes como tales —esto es, aparte de las soluciones de ecuaciones lineales—, aunque también los aplicó a la solución de ecuaciones lineales. Asimismo, proporcionó una regla para desarrollar un determinante usando menores de segundo orden y sus menores complementarios. En el sentido de que él fue quien más se concentró en los determinantes, se le considera el fundador de la teoría.

En un ensayo de 1772, «Recherches sur le calcul intégral et sur le système du monde» («Investigaciones sobre el cálculo integral y el sistema del mundo») [18], Laplace, refiriéndose al trabajo de Cramer y Bezout, probó algunas de las reglas de Vandermonde y generalizó su método de desarrollar determinantes al usar un conjunto de menores de r filas y los menores complementarios. Este método es aún conocido por su nombre [19].

La condición para que un conjunto de tres ecuaciones lineales no homogéneas con dos incógnitas tengan una solución común: la anulación del resultante expresa también el resultado de eliminar x e y de las tres ecuaciones; pero el problema de la eliminación se extiende en otras direcciones. Dados dos polinomios

$$f = a_0 x^m + \cdots + a_m$$
$$g = b_0 x^n + \cdots + b_n,$$

[16] *Hist. de l'Acad. des Sci., Paris*, 1764, 288-388.

[17] *Hist. de l'Acad. des Sci., Paris*, 1772, 516-532, pub. 1776.

[18] *Mém. de l'Acad. des Sci., Paris*, 1772, 267-376, pub. 1776 = *Œuvres*, 8, 365-406.

[19] Véase M. Bocher, *Introduction to Higher Algebra*, Dover (reimpresión), 1964, p. 26.

uno puede preguntarse por la condición para que $f=0$ y $g=0$ tengan una raíz común. Ya que la condición supone el hecho de que al menos un valor de x que satisface $f=0$ también satisface $g=0$, la sustitución de ese valor de x obtenido de $f=0$ y $g=0$ debe dar una condición en las a_i y b_i. Esta condición, o eliminante, o resultante, fue investigada primero por Newton. En su *Arithmetica Universalis (Aritmética Universal)* proporcionó reglas para la eliminación de x de las dos ecuaciones, las cuales podían ser de grado de dos a cuatro.

Euler, en el capítulo 19 del segundo volumen de su *Introductio*, da dos métodos de eliminación. El segundo es el precursor del método multiplicativo de Bezout, mejor descrito por Euler en su ensayo de 1764 [20]. El método de Bezout resultó ser el más ampliamente aceptado y, por tanto, debemos examinarlo. En su *Cours de mathématique* (*Curso de matemáticas*, 1764-69), considera las dos ecuaciones de grado n,

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0 \\ \phi(x) &= b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0 = 0. \end{aligned} \tag{10}$$

Se multiplica f por b_n y ϕ por a_n y se resta; después se multiplica f por $b_n x + b_{n-1}$ y ϕ por $a_n x + a_{n-1}$ y se resta; acto seguido f es multiplicada por $b_n x^2 + b_{n-1}x + b_{n-2}$ y ϕ por $a_n x^2 + a_{n-1}x + a_{n-2}$ y de nuevo se sustrae; y así sucesivamente. Cada una de las ecuaciones obtenidas de esta manera es de grado $n-1$ en x. Se puede considerar este conjunto de ecuaciones como un sistema de n ecuaciones lineales homogéneas con las incógnitas $x^{n-1}, x^{n-2}, \ldots, 1$. El resultante de este sistema de ecuaciones lineales, que es el determinante de los coeficientes de las incógnitas, es el resultante de las dos originales, $f=0$ y $\phi=0$. Bezout también proporcionó un método para encontrar la resultante cuando las dos ecuaciones no son del mismo grado [21].

La teoría de la eliminación también fue aplicada a dos ecuaciones, $f(x,y)=0$ y $g(x,y)=0$, de grado mayor que 1. La motivación era establecer el número de soluciones comunes a las dos ecuaciones o bien, geométricamente, encontrar el número de intersecciones de las curvas correspondientes a las ecuaciones. El sorprendente método

---

[20] *Mém. de l'Acad de Berlin*, 20, 1764, 91-104, pub. 1766 = *Opera*, (1) 6, 197-211.

[21] Se encuentra una exposición en W. S. Burnside y A. W. Panton, *The Theory of Equations*, Dover (reimpresión), 1960, vol. 2, p. 76.

para eliminar una incógnita de $f(x, y) = 0$ y $g(x, y) = 0$, primero esbozado por Bezout en su ensayo de 1764, lo presentó en su *Théorie générale des équations algébriques* (*Teoría general de las ecuaciones algebraicas*, 1779). La idea de Bezout era que multiplicando $f(x, y)$ y $g(x, y)$ por polinomios adecuados, $F(x)$ y $G(x)$, respectivamente, podía formar

$$R(y) = F(x)\, f(x, y) + G(x) g(x, y). \tag{11}$$

Más aún, buscó F y G tales que el grado de $R(y)$ resultara tan pequeño como fuera posible.

La cuestión del grado de la eliminante también fue contestada por Bezout en su *Théorie* (e independientemente por Euler en el ensayo de 1764). Ambos dieron como respuesta mn, el producto de los grados de f y g, y ambos probaron este teorema reduciendo el problema a uno de eliminación a partir de un conjunto auxiliar de ecuaciones lineales. Este producto es el número de puntos de intersección de las dos curvas algebraicas. Jacobi [22] y Minding [23] también dieron el método de Bezout de eliminación para las dos ecuaciones, pero ninguno menciona a Bezout. Pudo suceder que el trabajo de Bezout no les fuera conocido.

4. La teoría de números

En el siglo XVIII, la teoría de números aparecía como una serie de resultados desconectados. Los trabajos más importantes en la materia fueron el *Anleitung zur Algebra* (*Guía de álgebra*, edición alemana, 1770) de Euler y el *Essai sur la théorie des nombres* (*Ensayo sobre la teoría de números*, 1798) de Legendre. La segunda edición apareció en 1808 bajo el título *Théorie des nombres (Teoría de números)*, y una tercera edición difundida en tres volúmenes apareció en 1830. Los problemas y resultados que describiremos son un pequeño ejemplo de lo que se hizo.

En 1736, Euler demostró [24] el teorema menor de Fermat, a saber, que si p es un primo y a es primo con p, $a^p - a$ es divisible por p.

[22] *Journ. für Math.*, 15, 1836, 101-124 = *Gesam. Werke*, 3, 297-320.
[23] *Journ. für Math.*, 22, 1841, 178-183.
[24] *Comm. Acad. Sci. Petrop.*, 8, 1736, 141-146, pub. 1741 = *Opera*, (1), 2, 33-37.

Muchas pruebas de este mismo teorema fueron dadas por otros matemáticos de los siglos XVIII y XIX. En 1760, Euler generalizó este teorema [25] al introducir la función ϕ o indicador de n; $\phi(n)$ es el número de enteros $<n$ y primos con n, de tal forma que si n es primo, $\phi(n)$ es $n-1$. [La notación $\phi(n)$ fue introducida por Gauss.] Entonces Euler probó que si a es un primo relativo de n,

$$a^{\phi(n)} - 1$$

es divisible por n.

En cuanto a la famosa conjetura de Fermat sobre $x^n + y^n = z^n$, Euler probó [26] que era correcta para $n = 3$ y $n = 4$; el caso $n = 4$ había sido probado por Frenicle de Bessy. Este trabajo de Euler tuvo que ser completado por Lagrange, Legendre y Gauss. Legendre probó [27] entonces la conjetura para $n = 5$. Como veremos, la historia de los esfuerzos para probar la conjetura de Fermat es muy extensa.

Fermat también había conjeturado (cap. 13, sec. 7) que la fórmula

$$2^{2^n} + 1$$

proporcionaba números primos para un conjunto no especificado de valores de n. Esto es cierto para $n = 0$, 1, 2, 3 y 4. Sin embargo, Euler mostró en 1732 [28] que para $n = 5$ este número no es primo, uno de los factores es 641. De hecho, ahora es sabido que la fórmula no da primos para muchos otros valores de n, ya que ningún otro valor mayor de 4 ha sido encontrado para el que la fórmula dé un primo. Sin embargo, la fórmula es de interés en cuanto que reaparece en el trabajo de Gauss sobre la constructibilidad de polígonos regulares (cap. 31, sec. 2).

Un tema con muchas subdivisiones es el de la descomposición de enteros de varios tipos en otras clases de enteros. Fermat había afirmado que cada entero positivo es la suma de como mucho cuatro cuadrados (la repetición de un cuadrado como en $8 = 4 + 4$, está

---

[25] *Novi Comm. Acad. Sci. Petrop.*, 8, 1760/1, 74-104, pub. 1763 = *Opera*, (1), 3, 531-555.

[26] *Algebra*, parte II, segunda sección, 509-516 = *Opera*, (1), 1, 484-489 (para $n = 3$), y *Comm. Acad. Sci. Petrop.*, 10, 1738, 125-146, pub. 1747 = *Opera*, (1), 2, 38-59 (para $n = 4$).

[27] *Mém. de l'Acad. des Sci., Paris*, 6, 1823, 1-60, pub. 1827.

[28] *Comm. Acad. Sci. Petrop.*, 6, 1732/3, 103-107 = *Opera*, (1), 2, 1-5.

permitida si es contado el número de veces que aparece). A lo largo de un período de cuarenta años, Euler continuó intentando probar este teorema y obtuvo resultados parciales [29]. Usando parte del trabajo de Euler, Lagrange [30] probó el teorema. Ni Euler ni Lagrange obtuvieron el número de representaciones.

Euler, en un ensayo de 1754-55, únicamente se refirió a él y, en otro ensayo de la misma revista [31], probó la aseveración de Fermat de que cada primo de la forma $4n + 1$ es únicamente descomponible en suma de dos cuadrados. Sin embargo, la prueba de Euler no siguió el método de descenso que Fermat había bosquejado para este teorema. En otro ensayo [32], Euler probó también que cada divisor de la suma de los cuadrados de dos primos relativos es la suma de dos cuadrados.

Edward Waring (1734-98) estableció, en su *Meditationes Algebraicae* (*Meditaciones algebraicas*, 1770), el teorema ahora conocido como «Teorema de Waring»: que cada entero es, o un cubo o la suma de a lo más nueve cubos; también cada entero es o una potencia cuarta o la suma de a lo más 19 potencias cuartas. También conjeturó que cada entero positivo puede ser expresado como la suma de a lo más r potencias k-ésimas, con r dependiendo de k. Estos teoremas no los demostró [33].

Christian Goldbach, un prusiano enviado a Rusia, en una carta dirigida a Euler el 7 de junio de 1742, estableció sin prueba alguna que cada entero par es la suma de dos primos y cada entero impar, o bien es un primo o la suma de tres primos. La primera parte de la aserción es ahora conocida como la conjetura de Goldbach y sigue siendo un problema abierto. La segunda aserción se sigue de hecho de la primera, ya que si n es impar se substrae cualquier primo p de ella. Entonces $n - p$ es par.

Entre los resultados que tratan de la descomposición de números y son algo más especializados están la prueba de Euler de que $x^4 - y^4$

---

[29] *Novi Comm. Acad. Sci. Petrop.*, 5, 1754/5, 13-58, pub. 1760 = *Opera*, (1), 2, 338-372.

[30] *Nouv. Mém. de l'Acad de Berlin*, 1, 1770, 123-133, pub. 1772 = *Œuvres*, 3, 189-201.

[31] 5, 1754/5, 3-13, pub. 1760 = *Opera*, (1), 2, 328-337.

[32] *Novi Comm. Acad. Sci. Petrop.*, 4, 1752/53, 3-40, pub. 1758 = *Opera*, (1), 2, 295-327.

[33] El teorema general fue demostrado por David Hilbert (*Math. Ann.*, 67, 1909, pp. 281-300).

y $x^4 + y^4$ no pueden ser cuadrados [34]. Euler y Lagrange probaron muchas de las aserciones de Fermat en el sentido de que ciertos números primos pueden ser expresados en formas particulares. Por ejemplo, Euler demostró [35] que un primo de la forma $3n + 1$ puede ser expresado en la forma $x^2 + 3y^2$ de manera única.

Los números amigos y perfectos continuaron interesando a los matemáticos. Euler [36] proporcionó 62 pares de números amigos incluyendo los tres pares ya conocidos. Dos de sus pares eran incorrectos. También demostró, en un ensayo publicado póstumamente [37], el recíproco del teorema de Euclides: cada número par perfecto es de la forma $2^{p-1}(2^p - 1)$, donde el segundo factor es primo.

John Wilson (1741-1793), alumno sobresaliente en Cambridge, pero que se convirtió en abogado y juez, estableció un teorema que aún lleva su nombre: para cada primo p, la cantidad $(p - 1)! + 1$ es divisible por p; además, si es divisible por q, entonces q es primo. Waring publicó el argumento en sus *Meditationes Algebraicae (Mediationes Algebraicas)* y Lagrange lo probó en 1773 [38].

El problema de resolver la ecuación $x^2 - Ay^2 = 1$ con números enteros ya ha sido discutida (cap. 13, sec. 7); Euler, en un ensayo de 1732-1733, la llamó erróneamente la ecuación de Pell, y el nombre ha quedado. Su interés en esta ecuación se debía a que necesitaba sus soluciones para resolver $ax^2 + bx + c = y^2$ con enteros; escribió varios ensayos sobre este último tema. En 1759, Euler dio un método para resolver la ecuación de Pell al expresar $\sqrt{A}$ como una fracción continua [39]. La idea de Euler era que los valores x e y que satisfacen la ecuación son tales que las proporciones x/y son convergentes (en el sentido de fracciones continuas) a $\sqrt{A}$. Falló en probar que su método siempre proporciona soluciones y todas están dadas

[34] *Comm. Acad. Sci. Petrop.*, 10, 1738, pp. 125-146, pub. 1747 = *Opera* (1), 2, pp. 38-59; también en *Algèbre* (1770), Parte II, Cap. 13, sec. 202-8 = *Opera* (1), 1, pp. 436-443.

[35] *Novi Comm. Acad. Sci. Petrop.*, 8, 1760-1761, pp. 105-128, pub. 1763 = *Opera* (1), 2, pp. 556-575.

[36] «De numeris amicabilibus», *Opuscula varii argumenti*, 2, 1750, pp. 23-107 = *Opera* (1), 2, pp. 86-162.

[37] «De numeris amicabilibus», *Comm. Arith.*, 2, 1849, pp. 627-636 = *Opera Postuma*, 1, 1862, pp. 85-100 = *Opera* (1), 5, pp. 353-365.

[38] *Nouv. Mém. de l'Acad. de Berlin*, 2, 1771, pp. 125 y sgs., pub. 1773 = *Œuvres*, 3, pp. 425-438

[39] *Novi Comm. Acad. Sci. Petrop.*, 11, 1765, pp. 28-66, pub. 1767 = *Opera* (1), 3, pp. 73-111.

por las fracciones continuas que desarrollan $\sqrt{A}$. Lagrange[40] demostró en 1766 la existencia de soluciones para la ecuación de Pell y dio una forma más sencilla en ensayos posteriores[41].

Fermat afirmó que podía determinar cuándo la ecuación más general $x^2 - Ay^2 = B$ era resoluble con enteros y que podía resolverla cuando era posible. La ecuación fue resuelta por Lagrange en los dos ensayos mencionados.

El problema de proporcionar todas las soluciones enteras de la ecuación general

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0$$

donde los coeficientes son enteros fue atacado también. Euler proporcionó tipos incompletos de soluciones, y Lagrange[42] proporcionó la solución completa. En el siguiente volumen de sus *Memoires*[43] *(Memorias)* dio una solución más sencilla.

Tal vez el descubrimiento más original y de mayor trascendencia del siglo XVIII, respecto a la teoría de números, es la ley de reciprocidad cuadrática. Usa la noción de resultantes o residuos cuadráticos. En el lenguaje introducido por Euler en el ensayo de 1754-1755 y adoptado por Gauss, si existe una x tal que $x^2 - p$ es divisible por q, entonces se dice que p es el residuo o resto cuadrático de q; si no hay tal x, se dice que q no es residuo cuadrático de q. Legendre (1808) inventó un símbolo que ahora se usa para representar cualquiera situación: (p/q) y significa para cualquier número p y cualquier primo q

$$(p/q) = \begin{cases} 1 & \text{si } p \text{ es un residuo cuadrático de } q \\ -1 & \text{si } p \text{ no es un residuo cuadrático de } q. \end{cases}$$

También se entiende que $(p/q) = 0$ si p se divide de modo par en q.

La ley de reciprocidad cuadrática, en forma simbólica, establece

---

[40] *Misc. Taur.*, 4, 1766-1769, pp. 19 y sgs. = *Œuvres*, 1, pp. 671-731.

[41] *Mém. de l'Acad. de Berlin*, 23, 1767, pp. 165-310, pub. 1769, y 24, 1768, pp. 181-256, pub. 1770 = *Œuvres*, 2, pp. 377-535 y 655-726; también en las adiciones de Lagrange a su traducción del *Algebra* de Euler; véase la bibliografía.

[42] *Mém. de l'Acad. de Berlin*, 23, 1767, pp. 165-310, pub. 1769 = *Œuvres*, 2, pp. 377-535.

[43] 24, 1768, pp. 181-256, pub. 1770 = *Œuvres*, 2, pp. 655-726.

que si p y q son primos impares distintos, entonces

$$(p/q)(q/p) = (-1)^{(p-1)(q-1)/4}.$$

Esto significa que si el exponente de (-1) es par, p es un residuo cuadrático de q y q es el residuo cuadrático de p, o ninguno es un residuo cuadrático del otro. Cuando el exponente es impar, lo que sucede cuando p y q son de la forma $4k + 3$, un primo será residuo cuadrático del otro pero no al revés.

Detallamos la historia de esta ley. Euler, en un ensayo de 1783 [44], expuso cuatro teoremas y un quinto teorema resumen que establece muy claramente la ley de la reciprocidad cuadrática. Sin embargo, no demostró estos teoremas. El trabajo de este ensayo data de 1772 y aún incorpora trabajo previo. Kronecker observó en 1875 [45] que el enunciado de la ley está de hecho contenido en un ensayo muy anterior de Euler [46]. Sin embargo, la «prueba» de Euler estaba basada sobre cálculos. En 1785, Lagrange anunció la ley independientemente, aunque cita otro ensayo de Euler, en el mismo volumen del *Opuscula*, en su propio ensayo sobre este tema. Su prueba [47] era incompleta. En su *Théorie des nombres* [48] *(Teoría de números)*, enunció de nuevo la ley y dio una nueva prueba; sin embargo, ésta también era incompleta, porque supuso que hay un número infinito de primos en ciertas progresiones aritméticas. El deseo por encontrar lo que está detrás de esta ley y derivar sus muchas implicaciones constituyó un tema clave de investigaciones desde 1800 y ha llevado a resultados importantes, algunos de los cuales consideramos en capítulos posteriores.

El trabajo sobre la teoría de números en el siglo XVIII se cierra con el trabajo clásico de Legendre (*Théorie*, de 1798). A pesar de que contiene un buen número de resultados interesantes, en su dominio como en otros (tal como las integrales elípticas), Legendre no incluyó grandes innovaciones. Se podría reprocharle el presentar una colección de proposiciones de dónde sus concepciones generales po-

---

[44] *Opuscula Analytica*, 1, 1783, pp. 64-84 = *Opera* (1), 3, pp. 497-512.

[45] *Werke*, 2, pp. 3-10.

[46] *Comm. Acad. Sci. Petrop.*, 14, 1744-1746, pp. 151-181, pub. 1751 = *Opera* (1), 2, pp. 194-222.

[47] *Hist. de l'Acad. des Sci., Paris*, 1785, pp. 465-559, pub. 1788.

[48] 1798, pp. 214-226; 2da. ed., 1808, pp. 198-207.

drían haber sido extraídas, pero no lo fueron: lo hicieron más adelante sus sucesores.

Bibliografía

Cajori, Florian: «Historical note on the graphical representation of imaginaries before the time of Wessel», *Amer. Math. Monthly*, 19, 1912, pp. 167-171.

Cantor, Moritz: *Vorlesungen uber Geschichte der Mathematik*, B. G. Teubner, 1898 y 1924; Johnson Reprint Corp., 1965, vol. 3, cap. 107; vol. 4, pp. 153-198.

Dickson, Leonard E.: *History of the Theory of Numbers*, 3 vols., Chelsea (reimpresión), 1951.

— : «Fermat's last theorem», *Annals of Math.*, 18, 1971, pp. 161-187

Euler, Leonhard: *Opera Omnia* (1), vols. 1-5, Orell Fussli, 1911-1944.

— : *Vollständige Anleitung zur Algebra* (1770) = *Opera Omnia* (1), 1.

Fuss, Paul H. von, ed.: *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle*, 2 vols. (1843), Johnson Reprint Corp., 1967.

Gauss, Carl Friedrich: *Werke, Königliche Gessellschaft der Wissenschaften zu Göttingen*, 1876, vol. 3, pp. 3-121

Gerhardt, C. I.: *Der Briefwechsel von Gottfried Wilhelm Leibniz mit Mathematikern*, Mayer und Muller, 1899; Georg Olms (reimpresión), 1962.

Heath, Thomas L.: *Diophantus of Alexandria*, 1910, Dover (reimpresión), 1964, pp. 267-380.

Jones, P. S.: «Complex numbers: an example of recurring themes in the developments of mathematics», *The Mathematics Teacher*, 47, 1954, pp. 106-114, 257-263, 340-345.

Lagrange, Joseph Louis: *Œuvres*, Gauthier-Villars, 1867-1869, vols. 1-3, ensayos relevantes.

— : «Additions aux éléments d'algèbre d'Euler», *Œuvres*, Gauthier-Villars, 1877, vol. 7, pp. 5-179.

Legendre, Adrien Marie: *Théorie des nombres*, 4.ª ed., 2 vols., A Blanchard (reimpresión), 1955.

Muir, Thomas: *The theory of determinants in the historical order of development*, 1906, Dover (reimpresión), 1960, vol. 1, pp. 1-52.

Ore, Oystein: *Number theory and its History*, McGraw-Hill, 1948.

Pierpont, James: «Lagrange's place in the theory of substitutions», *Amer. Math. Soc. Bulletin*, 1, 1894-1895, pp. 196-204.

— : «Zur Geschichte der Gleichung des V. Grades (bis 1958)», *Monatshefte für Mathematik und Physik*, 6, 1895, pp. 15-68.

Smith, H. J. S.: *Report on the Theory of Numbers,* 1867, Chelsea (reimpresión), 1965; también en vol. 2 de *Collected mathematical papers of H. J. S. Smith,* 1894, Chelsea (reimpresión), 1965.

Smith, David Eugene: *A Source Book in Mathematics,* 1929, Dover (reimpresión) 1959, vol. 1, selecciones relevantes. Una de las selecciones es una traducción al inglés de la segunda demostración de Gauss del teorema fundamental del álgebra.

Struik, D. J.: *A Source Book in Mathematics, 1200-1800,* Harvard University Press, 1969, pp. 26-54, 99-122.

Vandiver, H. S.: «Fermat's last theorem», *Amer. Math. Monthly,* 53, 1946, pp. 555-578.

Whiteside, Derek T.: *The mathematical works of Isaac Newton,* Johnson Reprint Corp., 1967, vol. 2, pp. 3-134. Esta sección contiene la *Universal Arithmetic* (Aritmética Universal) en inglés.

Wussing, H. L.: *Die Genesis der abstrakten Gruppenbegriffes,* VEB Deutscher Verlag der Wissenschaften, 1969.

Capítulo 26
LAS MATEMATICAS DE 1800

> Cuando no podemos usar el compás de las matemáticas o la antorcha de la experiencia (...) es cierto que no podemos dar un solo paso hacia adelante.
>
> VOLTAIRE

1. La aparición del análisis

Si el siglo XVII ha sido correctamente llamado el siglo de los genios, entonces el siglo XVIII puede denominarse el siglo de los ingeniosos. A pesar de que los dos siglos fueron prolíficos, los autores del siglo XVIII, sin haber introducido ningún concepto tan original o fundamental como el del cálculo, sino mediante el ejercicio virtuoso de una técnica, explotaron y adelantaron el poder del cálculo para producir lo que ahora son ramas importantes: series, ecuaciones diferenciales ordinarias y parciales, geometría diferencial y el cálculo de variaciones. Al extender el cálculo a estas áreas diversas, construyeron lo que es ahora el más amplio dominio de las matemáticas, que nosotros llamamos análisis (a pesar de que la palabra ahora incluye un par de ramas las cuales los matemáticos del siglo XVIII apenas tocaron). El progreso en geometría analítica y álgebra, por otro lado, difícilmente fue algo más que una extensión menor de lo que el siglo XVII había iniciado. Aún el gran problema del álgebra, la solución de ecuaciones de grado n-ésimo, recibió atención porque era necesario en el análisis como, por ejemplo, en la integración por el método de fracciones simples.

Durante el primer tercio, o algo así, del siglo, los métodos geométricos eran usados con libertad; pero Euler y Lagrange, en particular, reconocieron la efectividad superior de los métodos analíticos, reemplazando deliberada y gradualmente los argumentos geométricos por analíticos. Los muchos libros de texto de Euler son ejemplos de cómo podía ser usado el análisis. Hacia el final del siglo, Monge revivió la geometría pura, aunque la utilizó en gran medida para dar un significado intuitivo y una guía al trabajo en análisis. Normalmente Monge es considerado como un geómetra, pero ello se debe a que, al trabajar durante una época en la que la geometría se había agotado, le infundió nueva vida mostrando su importancia, al menos para los fines arriba propuestos. De hecho, en un ensayo publicado en 1786 adjudicó implícitamente mayor importancia al análisis, haciendo la observación de que la geometría podía progresar, ya que el análisis era susceptible de ser empleado para estudiarla. Euler, como muchos otros, no buscó en principio nuevas ideas geométricas. El interés primario y los últimos resultados se dieron en el trabajo analítico.

El argumento clásico sobre la importancia del análisis se le debe a Lagrange en su *Mecánique analytique*. Escribe en su prefacio:

> Nosotros ya tenemos varios tratados sobre mecánica, pero el plan de éste es enteramente nuevo. Me he propuesto el problema de reducir esta ciencia [mecánica], y el arte de resolver problemas pertenecientes a ella, a fórmulas generales cuyo simple desarrollo proporciona todas las ecuaciones necesarias para la solución de cada problema (...). No se encontrarán diagramas en este trabajo. Los métodos que expongo en él no demandan ni construcciones ni razonamientos geométricos o mecánicos, sino únicamente operaciones algebraicas analíticas sujetas a un procedimiento uniforme y regular. Aquellos que gustan del análisis disfrutarán al ver cómo la mecánica se convierte en nueva rama de él, y me estarán agradecidos por haber extendido su dominio.

Laplace también subrayó el poder del análisis, en su *Exposition du système du monde (Exposición del sistema del mundo)* dice

> El análisis algebraico nos hace olvidar rápidamente el objetivo principal [de nuestras investigaciones] al enfocar nuestra atención en combinaciones abstractas y es únicamente al final cuando regresamos a nuestro objetivo original. Pero al abandonarse a las operaciones del análisis, se es llevado por la generalidad de su método y a las ventajas inestimables de transformar el

razonamiento por procedimientos mecánicos a resultados frecuentemente innacesibles a la geometría. Tal es la fecundidad del análisis, que es suficiente con traducir dentro de este lenguaje universal verdades particulares para ver surgir de su propia expresión multitud de nuevas e inesperadas verdades. Ningún otro lenguaje tiene la capacidad para la elegancia que surge de una larga sucesión de expresiones ligadas una a otra viniendo todas de una idea fundamental. Por lo tanto, los geómetras [matemáticos] de este siglo, convencidos de la superioridad del análisis, se han dedicado principalmente a extender su dominio y empujar hacia atrás sus fronteras [1].

Unas cuantas características del análisis requieren aclaración. El énfasis de Newton sobre la derivada y la antidiferenciación fue conservado, de tal manera que el concepto de sumación rara vez fue usado. Por otro lado, el *concepto* de Leibniz, esto es, la forma diferencial de la derivada y la notación de Leibniz se hicieron habituales, a pesar de que, a lo largo del siglo, las diferenciales leibnizianas no tuvieran un significado preciso. Las primeras diferenciales dy y dx de una función $y = f(x)$ fueron legitimizadas en el siglo XIX (cap. 40, sec. 3), pero las diferenciales de orden superior —aquellas usadas libremente por los hombres del siglo XVIII— no han sido asentadas sobre fundamentos establecidos con rigor aún hoy día. El siglo XVIII también continuó la tradición leibniziana de la manipulación formal de las expresiones analíticas y virtualmente acentuaron esta práctica.

La importancia asignada al análisis tuvo implicaciones que los autores del siglo XVIII no apreciaron. Profundizó la separación entre el concepto de número y la geometría e implícitamente menospreció el problema de los fundamentos mismos del sistema numérico, del álgebra y del propio análisis. Este problema sería crítico en el siglo XIX. Es curioso que los matemáticos del siglo XVIII aún se llamaban a sí mismos geómetras, término en boga durante las épocas precedentes en que la geometría preponderaba sobre las matemáticas.

2. La motivación del trabajo del siglo XVIII

Mucho más que en cualquier otro siglo, el trabajo matemático del siglo XVIII estuvo directamente inspirado por problemas de la física. Bien puede decirse que el objetivo del trabajo no eran las

---

[1] Libro V, Cap. 5 - *Œuvres*, 6, pp. 465-466.

matemáticas, sino la solución de problemas físicos; las matemáticas sólo significaban un medio para fines físicos. Laplace, aunque sea tal vez un caso extremo, las consideraba únicamente una herramienta al servicio de la física y de hecho él mismo estaba interesado por completo en su valor para la astronomía.

La rama principal de la física era, por supuesto, la mecánica, y particularmente la mecánica celeste. La mecánica se convirtió en el paraíso de las matemáticas, como Leonardo había predicho, porque sugirió muchas direcciones de investigación. Tan extenso era el interés de las matemáticas para los problemas de la mecánica que d'Alembert en su *Encyclopédie* y Denis Diderot (1713-1784) en sus *Pensées sur l'interprétation de la nature (Pensamientos sobre la interpretación de la naturaleza)* escribieron sobre una transición de la era de las matemáticas del siglo XVII a la era de la mecánica del siglo XVIII. En efecto, creían que la obra matemática de hombres como Descartes, Pascal y Newton era el *passé*, y que la mecánica llegaría a ser el máximo interés de los matemáticos. Generalmente, en su opinión, la matemática era útil tan sólo porque servía a la física. El siglo XVIII se concentró en torno a la mecánica de sistemas discretos de masas y de los medios continuos. Por un tiempo, la óptica estuvo relegada al fondo.

En realidad, acontecía lo opuesto a las afirmaciones sostenidas por Diderot y D'Alembert. La interpretación verdadera fue expresada —a la luz de desarrollos futuros— por Lagrange, cuando dijo que aquellos que gustaban del análisis estarían ya satisfechos de ver cómo la mecánica se convertía en una nueva rama de éste. Dicho más abiertamente, el programa iniciado por Galileo y seguido conspicuamente por Newton, orientado a expresar los principios físicos básicos como argumentos matemáticos cuantitativos y deducir nuevos resultados físicos mediante argumentos matemáticos, había avanzado mucho. A partir de ese hecho, la física se tornaba más y más matemática, al menos en las áreas donde los principios físicos eran suficientemente entendidos. La incorporación cada vez mayor de importantes ramas de la física dentro de marcos de referencia matemáticos fundó la física matemática.

Las matemáticas no empezaron únicamente a abarcar las ciencias, sino que, como no había una separación tajante entre ciencia y lo que llamaríamos ingeniería, los matemáticos se propusieron los problemas tecnológicos como una materia de estudio. Euler, por ejemplo, trabajó en el diseño de barcos, la acción de las velas, balística,

cartografía y otros problemas prácticos. Monge realizó su trabajo en excavación y relleno y el diseño de aspas de molino, tan en serio como cualquier problema de geometría diferencial o ecuaciones diferenciales.

J. F. (Jean-Etienne) Montucla (1725-1799), en su *Histoire des mathématiques (Historia de las Matemáticas,* 2.ª ed., 1799-1802) dividió las matemáticas en dos partes, una, «comprendiendo aquellas cosas que son puras y abstractas, y la otra, aquellas que uno llama compuestas o más ordinariamente, físico-matemáticas». Su segunda parte comprendía campos que pueden ser enfocados y tratados matemáticamente, tales como mecánica, óptica, astronomía, arquitectura militar y civil, acústica, música, así como los seguros. En la óptica incluyó la dióptrica y aun la optometría, catóptrica y perspectiva. La mecánica incluía la dinámica y la estática, hidrodinámica e hidrostática. La astronomía cubría geografía, astronomía teórica, astronomía esférica, gnómica (*e.g.*, cuadrantes de sol), cronología y navegación. Montucla también incluyó la astrología, la construcción de observatorios y el diseño de barcos.

3. El problema de la demostración

La mezcla de las matemáticas y la física constituyó un hecho crucial en el siglo XVIII, más que la motivación generada por la propensión a resolver los problemas físicos mediante las matemáticas. El principal avance, como se anotó, fue el análisis. Sin embargo, la propia base del cálculo no sólo resultaba poco clara, sino que había estado en duda casi desde el inicio de las investigaciones en la materia en el siglo XVII. El pensamiento del siglo XVIII era ciertamente libre e intuitivo. Se ignoraron cuestiones delicadas del análisis, tales como la convergencia de series e integrales, el intercambio del orden de la diferenciación y la integración, el uso de diferenciales de grado superior y los problemas en torno a la existencia de integrales y las soluciones de ecuaciones diferenciales. El que los matemáticos hayan sido capaces de avanzar de alguna manera se debió a que las reglas para efectuar las operaciones eran claras. Habiendo formulado los problemas físicos matemáticamente, los virtuosos se dedicaron a trabajar, y surgieron nuevas metodologías y conclusiones. Las matemáticas en sí mismas eran puramente formales. Euler estaba tan fascinado por las fórmulas que difícilmente podía abstenerse de manipu-

larlas. ¿Cómo se habían atrevido los matemáticos a aplicar sólo reglas y luego aseverar la seguridad de sus conclusiones?

El significado físico de las matemáticas guió los pasos matemáticos y aportó continuamente argumentos parciales para cubrir los pasos no matemáticos. En esencia, el razonamiento no se diferenciaba de una prueba de un teorema de geometría, donde algunos factores totalmente obvios en la figura son usados aunque ningún axioma o teorema los apoye. Finalmente, lo correcto de las conclusiones físicas ofreció seguridad de que las matemáticas debían ser correctas.

Otro elemento en el pensamiento de las matemáticas del siglo XVIII apoyaba los argumentos: el hombre confiaba en los símbolos mucho más que en su lógica. Ya que las series infinitas tenían la misma forma simbólica para todos los valores de x, la distinción entre los valores de x para los cuales la serie convergía y los valores para los que divergía no parecía demandar atención. Y, aunque reconocían que algunas series, tales como $1 + 2 + 3 + ...$, tenían una suma infinita, intentaron darle un significado a la suma en vez de cuestionar la sumatoria.

Del mismo modo, el uso un tanto libre de los números complejos descansaba sobre la confianza otorgada a los símbolos. A partir del hecho de que expresiones de segundo grado, $ax^2 + bx + c$, podían ser expresadas como producto de factores lineales donde los ceros eran reales, resultaba igualmente claro que debía haber factores lineales cuando los ceros eran complejos. Las operaciones formales del cálculo, diferenciación y antidiferenciación se aplicaron a nuevas funciones a pesar del hecho de que los matemáticos eran conscientes de la falta de claridad en sus ideas. Esta confianza en el formalismo los cegó de alguna manera. Así, su dificultad en ampliar la noción de función fue causada por el apego que tenían a la convicción de que las funciones debían ser expresables por medio de fórmulas.

Los matemáticos del siglo XVIII eran conscientes de los requerimientos matemáticos de prueba. Hemos visto que Euler intentó justificar el uso de sus series divergentes, y Lagrange, entre otros, ofreció una fundamentación para el cálculo. Pero los pocos esfuerzos para lograr rigor, significativos porque muestran que los niveles de rigor varían con el tiempo, no hicieron lógico el trabajo del siglo, y las gentes adoptaron casi de buena gana la postura de lo que no puede ser curado debe ser duradero. Estaban tan intoxicados con su éxito con los argumentos relacionados con la física, que cada vez se mostraban más indiferentes al rigor que faltaba. Impresiona mucho

la confianza extrema en las conclusiones, mientras que la teoría estaba tan mal apoyada. Dado que los matemáticos del siglo XVIII se habían propuesto seguir adelante tan ciegamente sin apoyo lógico, a este período se le ha llamado la época heroica de las matemáticas.

Quizá porque los pocos esfuerzos por hacer más sólido el análisis culminaron en fracasos y cuestiones adicionales de rigor surgidas del trabajo analítico subsecuente estaban desesperadamente faltas de solución, algunos matemáticos abandonaron su empeño en afianzar la teoría, y, como la zorra con las uvas, se alejaron conscientemente del rigor de los griegos. Sylvestre-François Lacroix (1765-1843), en la segunda edición de sus tres volúmenes *Traité du calcul différentiel et du calcul intégral (Tratado de cálculo diferencial y cálculo integral)*, dice, en el prefacio al volumen 1 (p. 11), «tales pequeñeces como aquellas por las que se preocupan los griegos, nosotros no las necesitamos». La actitud típica del siglo era: ¿Por qué complicarse probando con razonamientos oscuros cosas de las que uno nunca duda en primera instancia, o por demostrar lo que es más evidente por medio de lo que es menos evidente?

Hasta la geometría euclídea tuvo críticas sobre la base de que ofrecía pruebas donde ninguna parecía necesitarse. Clairaut afirmó en sus *Eléments de géométrie (Elementos de Geometría*, 1741):

> No es sorprendente que Euclides tenga dificultades para demostrar que dos círculos que se cortan no tienen el mismo centro, que la suma de los lados de un triángulo comprendido dentro de otro es más pequeña que la suma de los lados del triángulo menor. Este geómetra tuvo que convencer a sofistas obstinados que se glorificaban en rechazar las verdades más evidentes; de tal forma que la geometría debe, como la lógica, basarse en un razonamiento formal para refutar las sutilezas (...). Pero las cosas han cambiado. Todo razonamiento concerniente a lo que el sentido común sabe de antemano, sirve únicamente para esconder la verdad y agotar al lector, y es ignorado hoy día.

Esta actitud del siglo XVIII fue también expresada por Josef Maria Hoene-Wronski (1778-1853), quien fue un gran algoritmista pero no se interesaba por el rigor. La Comisión de la Academia de Ciencias de París criticó uno de sus ensayos porque carecía de rigor; Wronski respondió que ésto era «pedantería, que prefiere el medio al fin».

Los matemáticos eran, en su mayoría, conscientes de su propia indiferencia hacia el rigor. D'Alembert comentó en 1743: «Hasta el presente (...) más importancia se ha dado a engrandecer el edificio

que a iluminar la entrada, a elevarlo aún más que a asentar los cimientos». Por consiguiente, el siglo XVIII abrió nuevos caminos como hasta entonces no se había hecho. Las grandes creaciones, habida cuenta del pequeño número de matemáticos, fueron más numerosas que en cualquier otro siglo. Las gentes de los siglos XIX y XX, inclinados a despreciar el tosco, a menudo temerario, producto del siglo XVIII, subrayaron sus excesos y fallos a fin de enaltecer sus triunfos.

4. La base metafísica

A pesar de que los matemáticos sí reconocieron que sus creaciones no habían sido reformuladas en términos del método deductivo de Euclides, confiaban en la verdad de las matemáticas, apoyándose en parte —como hemos notado— en la veracidad física de sus conclusiones, pero también en bases filosóficas y teológicas. La verdad quedaba garantizada porque las matemáticas estaban desenterrando el orden matemático del universo. La clave de la filosofía de la última parte del siglo XVII y del XVIII, especialmente la expresada por Thomas Hobbes, John Locke y Leibniz, era la armonía preestablecida entre la razón y la naturaleza. Esta doctrina no había sido atacada desde el tiempo de los griegos. El problema era entonces: ¿carecían de la precisión de las pruebas puramente matemáticas las leyes matemáticas, que tan claramente se ajustan a la naturaleza? A pesar que el siglo XVIII únicamente desempolvó unas piezas, eran verdades fundamentales. La precisión asombrosa de las deducciones matemáticas, especialmente en mecánica celeste, significaba confirmación gloriosa de la confianza del siglo en el diseño matemático del universo.

Los autores del siglo XVIII estaban del mismo modo convencidos de que ciertos principios matemáticos debían ser ciertos por el diseño matemático que los incorporaba. Así, ya que un universo perfecto no permitiría el desperdicio, su acción era la mínima requerida para obtener sus propósitos. De aquí que el Principio de Acción Mínima fuese, como afirmaba Maupertuis y secundaba Euler en su *Methodus Inveniendi*, incuestionable.

La convicción de que el mundo estaba diseñado matemáticamente se derivó de la unión previa de la ciencia y la teología. Podemos recordar que la mayoría de las principales figuras de los siglos XVI y XVII no fueron sólo profundamente religiosos, sino que encontra-

ron en sus creencias teológicas inspiración y convicción vital para su trabajo científico. Copérnico y Kepler estaban convencidos de que la teoría heliocéntrica era cierta, porque Dios debió haber preferido una teoría más simple. Descartes creía que nuestras ideas innatas, entre ellas los axiomas de las matemáticas, son verdaderas, y que nuestro conocimiento es sólido, sobre la convicción de que Dios no nos engañaría, de tal forma que negar la verdad y claridad de las matemáticas sería negar a Dios. Newton, como hemos mencionado, concebía que el valor importante de sus esfuerzos científicos era el estudio del trabajo de Dios y el apoyo de la religión revelada. Muchos pasajes en sus trabajos son glorificaciones a Dios, y el escolio general, al final del libro III de sus *Principia,* en gran parte es un tributo a Dios, reminiscente de Kepler. La explicación de Leibniz sobre la concordancia entre los mundos real y matemático, y su defensa última de la aplicabilidad de su cálculo al mundo real, consistía en la unidad del mundo y Dios. Por tanto, las leyes de la realidad no podían desviarse de las leyes ideales de las matemáticas. El universo era lo más perfecto concebible, el mejor de todos los mundos posibles, y el pensamiento racional revelaba sus leyes.

A pesar de que permanecía la creencia en el diseño matemático de la naturaleza, el siglo XVIII eliminó finalmente las bases filosóficas y religiosas para tal creencia. El meollo de la posición filosófica básica, la doctrina de que el universo estaba diseñado por Dios, fue subestimada gradualmente por las explicaciones puramente físico-matemáticas. La motivación religiosa para el trabajo matemático había empezado a perder fuerza ya en el siglo XVII. Galileo había llamado la atención. Dice en una de sus cartas: «Aunque por mi parte, cualquier discusión sobre las Sagradas Escrituras puede esperar para siempre; ningún astrónomo o matemático que se mantenga dentro de los límites propios ha hecho incursiones en tales cosas». Así pues, Descartes, al afirmar la invariabilidad de las leyes de la naturaleza, había implícitamente limitado el papel de Dios. Newton confinó las acciones de Dios para mantener al mundo funcionando de acuerdo a su plan: usó la figura del lenguaje de un relojero vigilando un reloj en reparación. Así, el papel atribuido a Dios se hizo más y más restringido; y, en cuanto a la idea de que las leyes universales (que el propio Newton reveló) estaban incluyendo movimientos celestiales y terrestres, dichas concepciones empezaron a dominar la escena intelectual, y como continuó el acuerdo entre predicciones y observaciones, afirmando la perfección de estas leyes, entonces Dios se

hundió cada vez más en el fondo y las leyes matemáticas del universo se convirtieron en el foco de atención. Leibniz vio que los *Principia* implicaban un mundo funcionando de acuerdo con un plan con o sin Dios, y atacó el libro como anticristiano. Pero cuanto más se adelantaba en el desarrollo de las matemáticas en el siglo XVIII, más retrocedía la inspiración y motivación religiosa para el trabajo matemático.

Lagrange, a diferencia de Maupertuis y Euler, negó cualquier implicación metafísica en el Principio de Acción Mínima. La preocupación por obtener resultados significativos en física, reemplazó el interés por el diseño de Dios. Respecto a la preocupación de la física matemática, el total rechazo de Dios y de cualesquiera principios metafísicos que se apoyaban en su existencia, lo realizó Laplace. Hay una historia bastante bien conocida de cuando Laplace dio a Napoleón una copia de su *Mécanique celeste*, Napoleón indicó: «(...) M. Laplace, me dicen que ha escrito usted este extenso libro sobre el sistema del universo y nunca ha mencionado a su Creador». Se dice que Laplace repuso: «No he tenido necesidad de esa hipótesis». Hacia el final del siglo la etiqueta «metafísica» se aplicaba a un argumento en términos de reproche, aunque con frecuencia se usaba para condenar lo que un matemático no podía entender. Así, la teoría de las características de Monge, no entendida por sus contemporáneos, se consideró metafísica.

5. La expansión de la actividad matemática

Las academias de ciencias, fundadas en la mitad y final del siglo XVII, patrocinaron y apoyaron —a lo largo del siguiente siglo— la investigación matemática más que las universidades. Las academias también apoyaron a las revistas, convertidas en portadoras del nuevo trabajo. Tal vez el único cambio en las actividades de las academias sea que, en 1795, la Academia de Ciencias de París fue reconstituida como una de las tres divisiones del Instituto de Francia.

En Alemania, antes de 1800, las universidades no hacían investigación y ofrecían un par de años requeridos de artes liberales seguidos por especializaciones en leyes, teología o medicina. Los grandes matemáticos no se encontraban en las universidades, estaban ligados a la Academia de Ciencias de Berlín. Sin embargo, en 1810, Alexander von Humboldt (1769-1859) fundó la Universidad de Ber-

lín e introdujo la idea radical de que los profesores debían disertar sobre lo que a ellos les interesara y los estudiantes podían seguir los cursos que ellos quisieran. Así, por primera vez, los profesores podían disertar sobre sus investigaciones. Jacobi expuso en Koenigsberg, desde 1826 en adelante, su trabajo sobre funciones elípticas, aunque esto no era lo usual y otros maestros tenían que cubrir sus clases regulares. En el siglo XIX muchos de los reinos germanos, ducados y ciudades libres instauraron universidades que apoyaban la investigación hecha por los profesores.

Las universidades francesas del siglo XVIII, al menos hasta la Revolución, no eran mejores que las de Alemania. No obstante, el nuevo gobierno decidió fundar universidades de alto nivel para la enseñanza y la investigación. El trabajo organizativo estuvo a cargo de Nicolas de Condorcet, dedicado activamente a las matemáticas. La Ecole Polytechnique, fundada en 1794, tuvo en Monge y Lagrange a sus primeros profesores de matemáticas. Los estudiantes competían por ser admitidos y, supuestamente, debían recibir el entrenamiento que les permitiría convertirse en ingenieros u oficiales del ejército. De hecho, el nivel matemático de los cursos era muy alto y los graduados quedaban capacitados para realizar investigación matemática. A través de este entrenamiento y de las notas publicadas, la escuela ejercía una influencia amplia y fuerte. En 1808, el gobierno francés fundó la Ecole Normale Superieure, precedida por la Ecole Normale, fundada en 1794, y que únicamente duró unos meses. La nueva Ecole se dedicó a la formación de maestros y estuvo dividida en dos secciones: humanidades y ciencias. Aquí también los estudiantes competían por la admisión; eran ofrecidos cursos avanzados; las facilidades para estudiar e investigar eran grandes y a los mejores estudiantes se les orientaba hacia la investigación.

Las naciones europeas, durante el siglo XVIII, diferían considerablemente en su producción matemática. Francia era el país a la cabeza, seguido por Suiza. Alemania estaba relativamente inactiva, en cuanto a los científicos alemanes se entendía, aunque Euler y Lagrange gozaran del apoyo de la Academia de Ciencias de Berlín. Inglaterra también languidecía. Brook Taylor, Matthew Steward (1717-1785) y Colin Maclaurin eran los únicos matemáticos prominentes. Sorprende la escasa producción inglesa, a la vista de su gran actividad durante el siglo XVII, pero la explicación se encuentra rápidamente: los matemáticos ingleses no sólo se habían aislado personalmente del continente a consecuencia de la controversia entre

Newton y Leibniz, sino que también padecieron el seguir los métodos geométricos de Newton, dedicándose a estudiarlo en lugar de analizar la naturaleza. Incluso en su trabajo analítico usaron la notación de Newton para fluxiones y fluyentes y rechazaron leer cualquier cosa escrita en la notación de Leibniz. Más aún, en Oxford y Cambridge, a ningún judío o disidente de la Iglesia de Inglaterra le era dado, incluso, ser estudiante. En 1815, en Inglaterra, las matemáticas se agotaban en su última exhalación y la astronomía corría casi la misma suerte.

En el primer cuarto del siglo XIX, los matemáticos británicos empezaron a interesarse en el trabajo sobre el cálculo y sus extensiones, cuyo desarrollo en el continente había seguido sin cambios: en Cambridge (1813) fue formada la Sociedad Analítica con objeto de estudiar este trabajo. George Peacock (1791-1858), John Herschel (1792-1871), Charles Babbage y otros más se propusieron el análisis de los principios del «d-ismo» —esto es, la notación leibniziana en el cálculo, contra aquella de la «época del punto», o notación newtoniana—. Pronto, el cociente dy/dx reemplazó a $\dot{y}$, y los libros del texto y artículos del continente se hicieron accesibles a los estudiantes ingleses. Babbage, Peacock y Herschel tradujeron una edición en un solo volumen del *Traité (Tratado)* de Lacroix, publicándolo en 1816. En 1830, los ingleses estaban capacitados para unirse al trabajo del continente. El análisis en Inglaterra resultó ser en gran parte física matemática, aunque algunas vertientes del trabajo completamente nuevas, como la teoría algebraica de invariantes y la lógica simbólica, también fueron iniciadas en este país.

6. Un vistazo adelante

Como sabemos, hacia el final del siglo XVIII los matemáticos habían creado un buen número de nuevas ramas de las matemáticas. Pero los problemas de estas ramas se habían hecho extremadamente complicados, y, con muy pocas excepciones, no se habían creado nuevos métodos para tratarlos. Los matemáticos comenzaron a sentirse bloqueados. Lagrange escribió a D'Alembert el 21 de septiembre de 1781: «Me parece que la mina de las matemáticas es ya muy profunda, y que al menos que uno descubra nuevas vetas, será necesario, en algún momento, abandonarla. La física y la química ofrecen ahora una explotación más rica y fácil; también el gusto de

nuestro siglo parece ir por completo en esta dirección y no es imposible que las cátedras de geometría en la Academia se conviertan algún día en lo que las cátedras de árabe son actualmente en las universidades» [2]. Euler y D'Alembert estuvieron de acuerdo con Lagrange en que las matemáticas habían casi agotado sus ideas, y no vieron grandes mentes en el horizonte. Este miedo fue expresado tan temprano como en 1754 por Diderot en sus *Pensamientos sobre la interpretación de la Naturaleza:* «Me atrevo a decir que en menos de un siglo no quedarán tres grandes geómetras matemáticos en Europa. Esta ciencia muy pronto llegará a un punto estático donde los Bernoullis, Maupertuis, Clairuts, Fontaines, D'Alemberts y Lagranges la hubieran dejado. (...) No iremos más allá de este punto.»

Jean-Baptiste Delambre (1749-1822), secretario permanente de la sección de matemáticas y física del Instituto de Francia, en un informe (*Rapport historique sur les progrés des sciences mathématiques depuis 1789 et sur leur état actuel (Informe histórico sobre el progreso de las ciencias matemáticas desde 1789 y sobre su estado actual),* París, 1810) dijo: «Sería difícil y precipitado analizar las posibilidades que el futuro ofrece al avance de las matemáticas; en casi todas sus ramas se está bloqueando por barreras infranqueables; la perfección del detalle parece ser la única cosa que queda por hacer. Todas estas dificultades parecen anunciar que el poder de nuestro análisis se encuentra casi agotado (...).»

La predicción más sabia fue hecha en 1781 por Condorcet, quien estaba impresionado por el trabajo de Monge:

(...) a pesar de tantos trabajos coronados frecuentemente con el éxito, estamos lejos de haber agotado todas las aplicaciones del análisis a la geometría, y en lugar de creer que nos hemos acercado al final donde estas ciencias deben pararse porque ya han llegado al límite de las fuerzas del espíritu humano, nosotros debemos confesar que en su lugar estamos únicamente en los primeros pasos de una carrera inmensa. Estas nuevas aplicaciones [prácticas], independientemente de la utilidad que tengan en sí mismas, son necesarias para el progreso del análisis en general: dan nacimiento a cuestiones que uno no pensaría proponer; piden crear nuevos métodos. Los procesos técnicos son hijos de la necesidad; podríamos decir lo mismo de los métodos de las ciencias más abstractas. Pero nosotros debemos las últimas a necesidades más nobles, la necesidad de descubrir las verdades nuevas o de conocer mejor las leyes de la naturaleza.

---

[2] Lagrange, *Œuvres*, 13, p. 368.

Así, se ven en las ciencias muchas teorías brillantes que han permanecido sin aplicarse por mucho tiempo, convirtiéndose de repente en el fundamento de las más importantes aplicaciones; y, del mismo modo, aplicaciones que son muy simples en apariencia dando nacimiento a las ideas de las teorías más abstractas, para las cuales nadie ha sentido alguna necesidad, y dirigiendo el trabajo de geómetras [matemáticos] en estas direcciones (...).

Por supuesto Condorcet estaba en lo cierto. De hecho, en el siglo XIX, las matemáticas se expandieron aún más que en el siglo XVIII. En 1783, el año en que ambos, Euler y D'Alembert, murieron, Laplace tenía treinta y cuatro años de edad, Legendre treinta y uno, Fourier quince y Gauss seis.

La nueva matemática tal cual era la veremos en los siguientes capítulos. Pero debemos anotar aquí algunos de los agentes para la propagación de resultados a los que nos referiremos más tarde. Ante todo, hubo un vasto desarrollo en cuanto al número de revistas de investigación. El *Journal de l'Ecole Polytechnique* fue fundado al mismo tiempo que se organizó la escuela. En 1810, Joseph-Diez Gergonne (1771-1859) inició los *Annales de Mathématiques Pures et Appliquées,* que estuvo editándose hasta 1831, siendo ésta la primera revista puramente matemática. Al mismo tiempo, August Leopold Crelle (1780-1855), una figura notable porque fue un buen organizador y ayudó a un buen número de jóvenes a obtener trabajo en las universidades, inició el *Journal für die reine unangewandte Mathematik (Revista para matemáticas puras y aplicadas)* en 1826 [3]. Por medio de este título, Crelle intentó exponer que deseaba ensanchar la visión de los intereses de los matemáticos. A pesar de las intenciones de Crelle, la revista fue llenada rápidamente con artículos escritos por matemáticos especializados y con frecuencia se referían a esta publicación humorísticamente como el *Journal für reine unangewandte Mathematik (Revista para matemáticas puramente no aplicadas).* A esta revista también se la llama *Diario de Crelle* y, de 1855 a 1880, como *Borchardt's Journal.* En 1795 las *Mémoires de l'Académie des Sciences* se convirtieron en las *Mémoires de l'Académie des Sciences de l'Institut de France.* En 1835, la Academia de Ciencias de París también inició las *Comptes Rendus,* para cumplir el propósito de dedicar cuatro páginas o menos a noticias breves de nuevos resultados. Existe una historia apócrifa de que la restricción

[3] De aquí en adelante nos referiremos a él como *Jour. für Math.*

de las cuatro páginas intentaba limitar a Cauchy, quien había estado escribiendo prolíficamente.

En 1836, Liouville fundó el análogo francés del Diario de *Crelle*, El *Journal de Mathématiques Pures et Appliquées*, al que se refiere uno frecuentemente como *Diario de Liouville*. Louis Pasteur (1822-1895) fundó los *Annales Scientifiques de l'Ecole Normale Supérieure* en 1864, mientras que Gaston Darboux (1842-1917) lanzaba *Le Bulletin des Sciences Mathématiques* en 1870. Entre otras numerosas revistas, debemos mencionar los *Mathematische Annalen* (1868), el *Acta Mathematica* (1882) y el *American Journal of Mathematics*, la primera revista de matemáticas en los Estados Unidos de Norteamérica, fundada en 1878 por J. J. Sylvester, cuando era profesor en la Johns Hopkins University.

Otro tipo de agente ha promovido la actividad matemática desde el siglo XIX. Los matemáticos de diversos países formaron sociedades profesionales como la London Mathematical Society, primera en su género y organizada en 1865, la Societé Mathematique de France (1872), la American Mathematical Society (1888) y la Deutsche Mathematiker-Vereinigung (1890), las cuales mantuvieron reuniones regulares en las que se presentaban artículos; cada una patrocinaba una o más revistas, además de aquellas ya mencionadas, tales como el *Bulletin de la Societé Mathématique de France* y los *Proceedings of the London Mathematical Society*.

Como el bosquejo anterior de nuevas organizaciones y revistas indica, las matemáticas se expandieron enormemente en el siglo XIX y esto se hizo posible porque una pequeña aristocracia de matemáticos fue suplantada por un grupo mucho más amplio. El aumento de la enseñanza permitió la entrada de muchos más estudiantes de todos los niveles económicos. La tendencia se había iniciado ya en el siglo XVIII. Euler fue hijo de un pastor; D'Alembert fue un hijo ilegítimo criado por una familia de escasos recursos; Monge, hijo de un mercader, y Laplace nació en una familia campesina. La participación de las universidades en la investigación, la composición de libros de texto y la formación sistemática de científicos, iniciada por Napoleón, produjo mayor número de matemáticos.

Bibliografía

Boutroux, Pierre: *L'Ideal scientifique des mathématiques*, Libraire Felix Alcan, 1920.

Brunschvicg, Leon: *Les Etapes de la philosophie mathématique*, Presses Universitaires de France, 1947, Caps. 10-12.

Hankins, Thomas L.: *Jean d'Alembert: Science and Enlightenment*, Oxford University Press, 1970.

Hille, Einar: «Mathematics and mathematicians from Abel to Zermelo», *Mathematics Magazine*, 26, 1953, pp. 127-146.

Montucla, J. F.: *Histoire des mathématiques*, 2.ª ed., 4 vols., pp. 1799-1802, Albert Blanchard (reimpresión), 1960.

Capítulo 27
FUNCIONES DE UNA VARIABLE COMPLEJA

La trayectoria más corta entre dos verdades en
el dominio real pasa a través del dominio complejo.

JACQUES HADAMARD

1. Introducción

Desde el punto de vista técnico, la creación más original del siglo XIX fue la teoría de funciones de una variable compleja. Esta materia es con frecuencia llamada teoría de funciones; no obstante, su descripción abreviada implica más de lo que se intenta. Esta nueva rama de las matemáticas dominó el siglo XIX casi tanto como las extensiones directas del cálculo habían dominado el siglo XVIII. La teoría de funciones, una rama muy fértil de las matemáticas, ha sido llamada la alegría matemática del siglo. También ha sido aclamada como una de las teorías más armoniosas de las ciencias abstractas.

2. Los principios de la teoría de funciones complejas

Como ya hemos visto, los números complejos, y aun las funciones complejas, entraron en las matemáticas en relación con la integración de fracciones simples, la determinación de los logaritmos de números complejos y negativos, las aplicaciones conformes y la descomposición de un polinomio con coeficientes reales. De hecho, los

matemáticos del siglo XVIII hicieron muchas más cosas con los números y funciones complejos.

En su ensayo sobre hidrodinámica, *Ensayo sobre una nueva teoría de la resistencia de los fluidos* (1752), D'Alembert considera el movimiento de un cuerpo a través de un fluido homogéneo, ideal, carente de peso y en este estudio considera el siguiente problema. Busca determinar dos funciones p y q cuyas diferenciales son

$$dq = M\,dx + N\,dy, \qquad dp = N\,dx - M\,dy. \tag{1}$$

Ya que las cantidades N y M entran en ambas dp y dq, se sigue que

$$\frac{\partial p}{\partial x} = \frac{\partial q}{\partial y}, \qquad \frac{\partial p}{\partial y} = -\frac{\partial q}{\partial x}. \tag{2}$$

Estas ecuaciones son llamadas ahora las ecuaciones de Cauchy-Riemann. Las ecuaciones (2) dicen (cap. 19, sec. 6) que $q\,dx + p\,dy$ y $p\,dx - q\,dy$ son diferenciales exactas de ciertas funciones. De aquí que las expresiones (usaremos i para $\sqrt{-1}$, aunque esto solamente fue hecho ocasionalmente por Euler y hecho práctica común por Gauss)

$$q\,dx + p\,dy + i(p\,dx - q\,dy) = (q + ip)\left(dx + \frac{dy}{i}\right)$$

y

$$q\,dx + p\,dy - i(p\,dx - q\,dy) = (q - ip)\left(dx - \frac{dy}{i}\right)$$

son también diferenciales totales, y así $q + ip$ es una función de $x + y/i$ y $q - ip$ es una función de $x - y/i$. D'Alembert escribe

$$q + ip = \xi\left(x + \frac{y}{i}\right) + i\zeta\left(x + \frac{y}{i}\right) \tag{3}$$

$$q - ip = \xi\left(x - \frac{y}{i}\right) - i\zeta\left(x - \frac{y}{i}\right) \tag{4}$$

donde ξ y ζ son funciones que han de ser determinadas y que D'Alembert determina en casos especiales. Añadiendo y sustrayendo (3)

y (4) obtiene p y q. La significación de esto es que se ha mostrado que p y q son las partes real e imaginaria de una función compleja.

Euler mostró cómo usar funciones complejas para evaluar integrales reales. Escribió una serie de ensayos desde 1776 hasta su muerte en 1783, que fueron publicados a partir de 1788, y dos de los cuales fueron publicados en 1793 y 1797 [1]. Euler subraya que toda función de z para la cual $z = x + iy$ toma la forma $M + iN$, donde M y N son funciones reales, también toma, para $z = x - iy$, la forma $M - iN$. Esto, afirma, es el teorema fundamental de los números complejos. Usa esta aserción para evaluar integrales reales.

Supóngase

$$\int Z(z)\, dz = V \tag{5}$$

donde z es real. Escribe $z = x + iy$ de tal forma que V se convierte en $P + iQ$. Entonces

$$P + iQ = \int (M + iN)(dx + i\, dy), \tag{6}$$

donde $M + iN$ es ahora la forma compleja de $Z(z)$. Por medio de esta aserción básica,

$$P - iQ = \int (M - iN)(dx - i\, dy), \tag{7}$$

de tal manera que al separar las partes real e imaginaria,

$$P = \int M\, dx + N\, dy, \qquad Q = \int N\, dx + M\, dy. \tag{8}$$

Entonces $M\, dx - N\, dy$ y $N\, dx + M\, dy$ son las diferenciales exactas de P y Q, respectivamente, de lo cual se sigue que

$$\frac{\partial M}{\partial y} = -\frac{\partial N}{\partial x}, \qquad \frac{\partial N}{\partial y} = \frac{\partial M}{\partial x}. \tag{9}$$

[1] *Nova Acta Acad. Sci. Petrop.*, 7, 1789, 99-133, pub. 1793 = *Opera*, (1), 19, 1-44; *ibid.*, 10, 1792, 3-19, pub. 1797 = *Opera*, (1), 19, 268-286.

Entonces al sustituir $z = x + iy$ en $Z(z)$, «se obtienen dos funciones M y N que poseen la notable propiedad de que $\partial M/\partial y = -\partial N/\partial x$ y $\partial M/\partial x = \partial N/\partial y$; P y Q tienen propiedades similares». Aquí Euler subraya que M y N, las partes real e imaginaria de una función compleja, satisfacen las ecuaciones de Cauchy-Riemann. Sin embargo, el punto principal es usar las integrales (8) para calcular (5), ya que P es igual a la V original. Para reducir las integrales en (8) a integrales de funciones de una sola variable, Euler reemplaza $z = x + iy$ en (5) por $z = r(\cos\theta + i\,\mathrm{sen}\,\theta)$ y mantiene θ constante. Esto equivale a integrar a lo largo de un rayo que pasa por el origen del plano complejo. Más adelante usa este método para calcular algunas integrales.

Laplace también utilizó funciones complejas para evaluar integrales. En una serie de ensayos que empiezan en 1782 y que finalizan con su famosa *Théorie analytique des probabilités* (*Teoría analítica de las probabilidades*, 1812), pasa de integrales reales a complejas, como lo hizo Euler, para evaluar integrales reales. Laplace reclamó la prioridad, ya que los ensayos de Euler fueron publicados después que los suyos. Sin embargo, aún los ensayos de 1793 y 1797 ya mencionados se leyeron a la Academia de San Petersburgo en marzo de 1777. Incidentalmente, en este trabajo Laplace presentó lo que ahora nosotros llamamos el método de la transformada de Laplace para la solución de ecuaciones diferenciales.

La obra de Euler, D'Alembert y Laplace constituyó un progreso significativo en la teoría de funciones. Sin embargo, existe una limitación esencial en su trabajo. Ellos dependían de la separación de las partes real e imaginaria de $f(x + iy)$ para llevar a cabo su trabajo analítico. La función compleja no era realmente la entidad básica. Es claro que estos autores estaban aún muy insatisfechos con el uso de las funciones complejas. Laplace, en su libro de 1812, señala: «Esta transición de real a imaginario puede ser considerada como un método heurístico, que es como el método de inducción, usado por largo tiempo por los matemáticos. Sin embargo, si se usa el método con gran cuidado y limitación, uno siempre será capaz de probar los resultados obtenidos». Insiste en que los resultados deben ser verificados.

3. La representación geométrica de los números complejos

Un paso vital que hizo el surgimiento de una teoría de funciones de una variable compleja intuitivamente más razonable fue la representación geométrica de los números complejos y de las operaciones algebraicas entre estos números. Que todos aquellos hombres —Cotes, De Moivre, Euler y Vandermonde— pensaran realmente los números complejos como puntos en el plano se sigue del hecho que, al intentar resolver $x^n - 1 = 0$, pensaron las soluciones

$$\cos \frac{2k\pi}{n} + i \operatorname{sen} \frac{2k\pi}{n}$$

como los vértices de un polígono regular. Euler, por ejemplo, reemplazó x e y, visualizándolos geométricamente como un punto en el plano de coordenadas, por $x + iy$ y entonces lo representó por $r(\cos \theta + \mathrm{i} \operatorname{sen} \theta)$, los cuales a su vez representa como las coordenadas polares r y θ. De aquí que se pueda decir que la representación de los números complejos como coordenadas de puntos en el plano era conocida hacia 1800. Sin embargo, no fue hecha la identificación decisiva de los dos, ni las operaciones algebraicas con los números complejos tenían aún un significado geométrico. También faltaba la idea de trazar los valores de una función $u + iv$ de $x + iy$ como puntos de otro plano.

En 1797 el autodidacta navegante (nacido noruego) Caspar Wessel (1745-1818) escribió un ensayo titulado «Sobre la representación analítica de la dirección; un intento», el cual fue publicado en las memorias de 1799 de la Academia Real de Dinamarca. Wessel buscó representar geométricamente segmentos lineales dirigidos (vectores) y operaciones entre ellos. En este ensayo un eje de imaginarios con $\sqrt{-1}$ (él escribe ε por $\sqrt{-1}$) como una unidad asociada es introducido junto con el eje x usual con unidad real 1. En la representación geométrica de Wessel, el vector OP (Fig. 27.1) es el segmento lineal OP trazado desde el origen O en el plano de unidades $+1$ y $\sqrt{-1}$, y el vector es representado por el número complejo, $a + b\sqrt{-1}$. Análogamente, el vector OQ es el segmento lineal OQ y es representado por otro número, digamos $c + d\sqrt{-1}$.

Entonces Wessel define las operaciones con vectores al definir las operaciones con números complejos en términos geométricos. Sus definiciones de las cuatro operaciones son prácticamente las que

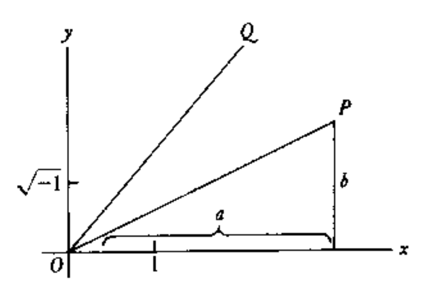


FIGURA 27.1

aprendemos hoy en día. Así, la suma de $a + bi$ y $c + di$ es la diagonal del paralelogramo determinado por los lados adyacentes OP y OQ. El producto de $a + bi$ y $c + di$ es un nuevo vector OR, digamos, tal que OR es a OQ como OP es a la unidad real y el ángulo formado por OR y el eje x es la suma de los ángulos formados por OP y OQ. Claramente, Wessel pensó en términos de vectores en lugar de asociar números complejos con puntos en el plano. Aplica su representación geométrica de los vectores a problemas de geometría y trigonometría. A pesar de su gran mérito, el ensayo de Wessel no atrajo la atención hasta 1897, cuando fue republicado en una traducción francesa.

Una interpretación geométrica de números complejos en cierta forma diferente fue proporcionada por el suizo Jean Robert Argand (1768-1822). Argand, que era autodidacta y tenedor de libros, publicó un libro breve, *Essai sur une manière de représenter les quantités imaginaires dans les constructions géométriques* (*Ensayo sobre una manera de representar las cantidades imaginarias en las construcciones geométricas*, 1806) [2]. Aquí observa que los números negativos eran una extensión de los números positivos que resulta de combinar la dirección con la magnitud. Entonces pregunta, ¿podemos extender el sistema numérico real al añadir nuevos conceptos? Consideremos la secuencia 1, x, -1. ¿Podemos encontrar una operación que convierta a 1 en x, y luego al repetirse sobre la x con-

[2] Un buen número de artículos sobre las ideas de Argand y las ideas de otros escritores sobre la representación geométrica de los números complejos se encuentran en el volumen 4 (1813-1814) y en el volumen 5 (1814-1815) de los *Annales des Mathématiques* de Gergonne.

vierta x en -1? Si rotamos OP (Fig. 27.2) en contra de las manecillas del reloj alrededor de O un ángulo de 90 ° y después repetimos esta rotación, vamos de P a Q al repetir la operación dos veces. Pero, nota Argand, esto es precisamente lo que sucede si multiplicamos 1 por $\sqrt{-1}$ y después multiplicamos este producto por $\sqrt{-1}$; esto es, obtenemos -1. De aquí que podamos pensar en $\sqrt{-1}$ como una rotación de 90° contra las manecillas del reloj, digamos, y $-\sqrt{-1}$ como una rotación (en dirección de las manecillas del reloj) de 90°.

Para utilizar ese significado operacional de los números complejos, Argand decidió que un segmento lineal típico OB (Fig. 27.3) emanando del origen, lo que él llama una línea directa, debe estar representado por $r(\cos\alpha + i\,\text{sen}\,\alpha)$, donde r es la longitud. También consideraba los números complejos $a + bi$ como simbolizando la combinación geométrica OB de a y bi. Argand, como Wessel, mostró cómo los números complejos pueden ser geométricamente sumados y multiplicados y aplicó estas ideas geométricas a probar teoremas de trigonometría, geometría y álgebra. A pesar de que el libro de Argand despertó cierta controversia acerca de la interpretación geométrica de los números complejos, esta fue la única contribución a las matemáticas que hizo y su trabajo tuvo poco impacto. Nosotros, sin embargo, aún hablamos del diagrama de Argand.

Gauss fue más eficaz en lograr la aceptación de los números complejos. Los usó en varias de sus pruebas del teorema fundamental del álgebra (cap. 25, sec. 2). En las primeras tres pruebas (1799, 1815 y 1816) presupone la correspondencia uno a uno de puntos del plano cartesiano con los números complejos. Esto no es de hecho el dibujo de $x + iy$, sino el de x e y como coordenadas de un punto en el plano real. Más aún, las demostraciones en realidad no usan la teoría de funciones complejas porque separa las partes real e imagi-

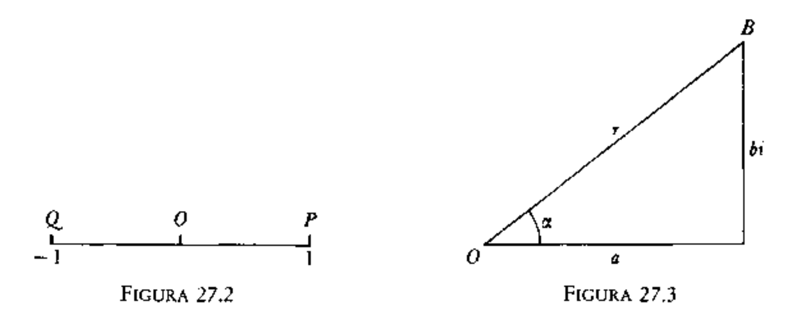


FIGURA 27.2

FIGURA 27.3

naria de las funciones. Gauss es más explícito en una carta a Bessel de 1811 [3], donde dice que $a + bi$ está representada por el punto (a,b) y que puede irse de un punto a otro del plano complejo por muchas trayectorias. No hay duda, si uno juzga por el pensamiento mostrado en las tres pruebas y en otros trabajos no publicados, algunos de los cuales discutiremos en breve, que para 1815 Gauss estaba en completa posesión de la teoría geométrica de números y funciones complejas, a pesar de que dijo en una carta de 1825 que «la verdadera metafísica de $\sqrt{-1}$ es elusiva».

Sin embargo, si Gauss aún tenía algunos escrúpulos, para 1831 ya los había superado, y describió públicamente la representación geométrica de los números complejos. En el segundo comentario a su ensayo «Theoria Residuorum Biquadraticorum» (Teoría de los Residuos Bicuadráticos) [4] y en el «Anzaige» (anuncio y breve descripción) de este ensayo que escribió para el *Göttingische gelehrte Anzeigen* del 23 de abril de 1831 [5], Gauss es muy explícito acerca de la representación geométrica de los números complejos. En el inciso 38 del ensayo no solamente proporciona la representación de $a + bi$ como un punto (no como un vector, como Wessel y Argand) en el plano complejo, sino que además describe la adición y multiplicación geométrica de los números complejos. En el «Anzeige» [6], Gauss menciona que la transferencia de la teoría de los residuos bicuadráticos en el dominio de los números complejos puede molestar a quienes no están familiarizados con estos números y que los puede dejar con la impresión que la teoría de los residuos queda en el aire. Por lo tanto, repite lo que ha dicho en el propio ensayo acerca de la representación geométrica de los números complejos. Señala que mientras las fracciones, números negativos y números reales eran ahora bien entendidos, los números complejos habían sido meramente tolerados a pesar de su gran valor. Para muchos, éstos habían aparecido para ser sólo un juego con los símbolos. Pero en esta representación geométrica uno encuentra que: «*el significado intuitivo de los números complejos está completamente establecido y no se necesita nada más para admitir estas cantidades en el dominio de la aritmética* (itálicas añadidas)». También afirma que si a las

[3] *Werke*, 8, 90-92.

[4] *Comm. Soc. Gott.*, 3, 1832 = *Werke*, 2, 95-148; el contenido principal de este artículo será discutido en el Capítulo 34, sec. 2.

[5] *Werke*, 2, 169-178.

[6] *Werke*, 2, 174 sgs.

unidades 1, -1 y $\sqrt{-1}$ no se les hubiera asignado los nombres de unidades positivas, negativas e imaginarias, sino que se les hubiera llamado directa, inversa y lateral, la gente no habría tenido la impresión de que existía algún misterio oscuro en estos números. La representación geométrica, afirma, sitúa la verdadera metafísica de los números imaginarios bajo una nueva luz. Introdujo el término «número complejo» como opuesto al de número imaginario [7], y usó i para $\sqrt{-1}$.

4. Los fundamentos de la teoría de funciones complejas

Gauss también introdujo algunas ideas básicas acerca de las funciones de una variable compleja. En una carta a Bessel de 1811 [8], acerca de un ensayo de Bessel sobre la integral logarítmica, $\int dx/x$, Gauss señala la necesidad de tener en cuenta los límites imaginarios (complejos). Enseguida pregunta, ¿qué debería uno interpretar por $\int f(x)\,dx$ [Gauss escribe $\int \phi x\,dx$] cuando el límite superior es $a + bi$? De hecho, si uno deseara aclarar los conceptos, uno debería suponer que x toma incrementos pequeños a partir de ese valor de x para el cual la integral es 0 hasta el valor $a + bi$ y entonces sumar todos los $\phi(x)\,dx$... Pero el paso continuo de un valor de x a otro en el plano complejo tiene lugar sobre una curva y es por lo tanto posible ir sobre muchas trayectorias. Afirmo ahora que la integral $\int \phi(x)\,dx$ tiene un valor único aún tomada sobre varias trayectorias siempre que $\phi(x)$ tome un valor único y no se haga infinita en el espacio comprendido por las dos trayectorias. Este es un teorema muy bello cuya demostración no es difícil y que proporcionaré en alguna ocasión pertinente». Gauss no dio esta prueba. También afirma que si $\phi(x)$ se hace infinita, entonces $\int \phi(x)\,dx$ puede tomar muchos valores, dependiendo de si uno toma una trayectoria cerrada que encierre el punto en el que $\phi(x)$ se haga infinita una, dos o muchas veces.

Más adelante, Gauss regresa al caso particular de $\int dx/x$ y dice que, saliendo a partir de $x = 1$ y dirigiéndose a algún valor $a + bi$, se obtiene un único valor para la integral si la trayectoria no encierra $x = 0$; pero si lo hace, se debe añadir $2\pi i$ o $-2\pi i$ al valor obtenido al ir de $x = 1$ a $x = a + bi$ sin encerrar $x = 0$. Así, existen muchos

[7] *Werke*, 2, p. 102.
[8] *Werke*, 8, 90-92.

logaritmos para una $a + bi$ dada. Más adelante, en la misma carta, Gauss dice que la investigación de las integrales de funciones de argumentos complejos debe proporcionar resultados muy interesantes. Así, aun antes de que Gauss hubiera publicado sus segunda y tercera pruebas del teorema fundamental —donde, como en la primera prueba, evitó el uso directo de números y funciones complejas excepto para escribir ocasionalmente $a + b\sqrt{-1}$—, ya tenía ideas muy definidas acerca de las funciones complejas y de las integrales de tales funciones.

Poisson notó en 1815, y lo discutió en un ensayo de 1820 [9], el uso de las integrales de funciones complejas tomadas sobre trayectorias en el plano complejo. Como ejemplo da

$$\int_{-1}^{1} \frac{dx}{x}. \tag{10}$$

Aquí pone $x = e^{i\theta}$, donde θ va desde $(2n + 1)\pi$ a 0, y obtiene, al tratar la integral como un límite de sumas, el valor $-(2n + 1)\pi i$.

Más adelante nota que el valor de la integral no tiene que ser el mismo cuando es tomada sobre una trayectoria imaginaria o cuando es tomada sobre una real. Menciona el ejemplo,

$$\int_{-\infty}^{\infty} \frac{\cos ax}{b^2 + x^2}\, dx, \tag{11}$$

donde a y b son constantes positivas. Hace $x = t + ik$, donde k es constante y positiva, y obtiene el valor $\pi(e^{-ab} - e^{ab})|2\pi$ para $k > b$ y πe^{-ab} *para* $k < b$. El segundo valor es también el correcto para $k = 0$. Entonces, para dos valores diferentes de k, lo que significa dos trayectorias diferentes, uno obtiene dos resultados diferentes. Poisson fue el primero en llevar a cabo integraciones sobre una trayectoria en el plano complejo.

A pesar de que estas observaciones de Gauss y Poisson son en realidad significativas, ninguno publicó algún ensayo importante en la teoría de funciones complejas. Esta teoría fue fundada por Augustin Louis Cauchy. Nacido en París en 1789, en 1805 ingresó en la Ecole Polytechnique, donde estudió ingeniería. Debido a su pobre salud le fue aconsejado por Lagrange y Laplace dedicarse a las ma-

9 *Jour. de l'Ecole Poly.*, 11, 1820, 295-341.

temáticas. Tuvo cátedra en la Ecole Polytechnique, la Sorbonne y el Collège de France. La política tuvo efectos inesperados en su carrera. Fue un ardiente monárquico y apoyó a los Borbones. Cuando en 1830 una rama distante de los Borbones tomó el poder en Francia, rehusó jurar lealtad a la nueva monarquía y renunció a su plaza en la Ecole Polytechnique. Se exilió en Turín, dedicándose a enseñar latín e italiano por algunos años. En 1838 regresó a París, donde trabajó como profesor de diversas instituciones religiosas, hasta el momento, después de la revolución de 1848, cuando el gobierno tomó su control y abolió los juramentos y lealtades. En 1848, Cauchy tuvo la cátedra de astronomía matemática en la Faculté des Sciences de la Sorbonne. A pesar de que Napoleón III restauró el juramento en 1852, le permitió a Cauchy olvidarlo. Al condescendiente gesto del Emperador, Cauchy respondió donando su salario a los pobres de Sceaux, donde vivía. Cauchy, un profesor admirable y uno de los más grandes matemáticos, murió en 1857.

Cauchy tenía intereses universales. Conocía la poesía de su época y fue el autor de un trabajo sobre la prosa hebrea. En matemáticas escribió más de setecientos artículos, siguiendo en número únicamente a Euler. En una edición moderna, sus trabajos llenan veintiséis volúmenes y abarcan todas las ramas de las matemáticas. En mecánica escribió importantes trabajos sobre el equilibrio de pesos y de membranas elásticas y sobre ondas en medios elásticos. En la teoría de la luz se ocupó de la teoría de ondas que Fresnel había empezado y de la dispersión y la polarización de la luz. Avanzó inmensamente en la teoría de los determinantes y contribuyó con teoremas básicos a las ecuaciones diferenciales ordinarias y parciales.

El primer artículo significativo de Cauchy en la dirección de la teoría de funciones complejas es su «Memoire sur la théorie des intégrales définies» («Memoria sobre la teoría de integrales definidas»). Este ensayo se leyó ante la Academia de París en 1814; sin embargo, no fue enviado para su publicación hasta 1825 y fue publicado en 1827 [10]. En la publicación Cauchy añadió dos notas que seguramente reflejan desarrollos entre 1814 y 1825 y la posible influencia del trabajo de Gauss durante este período. Por el momento, consideremos el propio artículo. Cauchy afirma en el prefacio que se orientaba en este trabajo a tratar de rigorizar el paso de real a imaginario en procedimientos usados por Euler desde 1759 y por

[10] *Mém. des sav. étrangers*, (2), 1, 1827, 599-799 = *Œuvres*, (1), 1, 319-506.

Laplace desde 1782 para evaluar integrales definidas, y de hecho, Cauchy cita a Laplace, quien observó que el método requería rigorización. Pero el propio artículo no trata este problema. Trata la cuestión del cambio del orden de la integración en integrales dobles que surgen en las investigaciones hidrodinámicas. Euler había dicho en 1770 [11] que este intercambio estaba permitido cuando los límites para cada una de las variables bajo el signo de la integral fueran independientes entre sí, y Laplace, aparentemente, estaba de acuerdo, porque utilizó este recurso en repetidas ocasiones.

Específicamente, Cauchy trató la relación

$$\int_{x_0}^{X}\int_{y_0}^{Y} f(x, y)\, dy\, dx = \int_{y_0}^{Y}\int_{x_0}^{X} f(x, y)\, dx\, dy, \qquad (12)$$

donde x_0, y_0, X e Y son constantes (fig. 27.4). Este cambio en el orden de integración es lícito cuando $f(x,y)$ es continua dentro y sobre el borde de la región. Luego introduce dos funciones $V(x,y)$ y $S(x,y)$ tales que

$$\frac{\partial V}{\partial y} = \frac{\partial S}{\partial x} \quad \text{y} \quad \frac{\partial V}{\partial x} = -\frac{\partial S}{\partial y}. \qquad (13)$$

En 1777, Euler ya había mostrado cómo obtener tales funciones (véase [5], [8] y [9]). Ahora, Cauchy considera una $f(x,y)$ que está dada por $\partial V/\partial y = \partial S/\partial x$. En (12) reemplaza la f del lado izquierdo

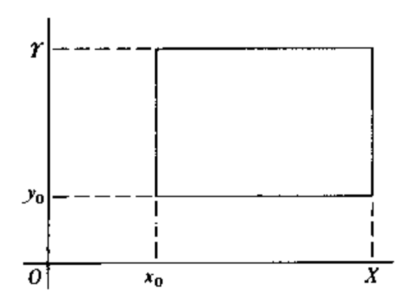


FIGURA 27.4

[11] *Novi Comm. Acad. Sci. Petrop.*, 14, 1769, 72-103, pub. 1770 = *Opera*, (1), 17, 289-315.

por $\partial V/\partial y$, y la $\int$ del lado derecho por $\partial S/\partial x$, de tal forma que

$$\int_{x_0}^{X}\int_{y_0}^{Y}\frac{\partial V}{\partial y}\,dy\,dx=\int_{y_0}^{Y}\int_{x_0}^{X}\frac{\partial S}{\partial x}\,dx\,dy, \qquad (14)$$

y usando la segunda ecuación en (13) obtiene

$$\int_{x_0}^{X}\int_{y_0}^{Y}\frac{\partial S}{\partial y}\,dy\,dx=-\int_{y_0}^{Y}\int_{x_0}^{X}\frac{\partial V}{\partial x}\,dx\,dy. \qquad (15)$$

Dichas igualdades pueden ser usadas para evaluar integrales dobles en cualquier orden de integración; sin embargo, no hacen intervenir funciones complejas. Cuando Cauchy afirma en su *Introducción* [12] que «establecerá rigurosa y directamente el paso de real a imaginario (complejo)», son las ecuaciones (13) las que tiene en su mente. Cauchy afirma [13] que estas dos ecuaciones contienen la teoría completa del paso de real a imaginario.

Todo lo anterior se encuentra en el artículo de 1814, y realmente no da indicaciones explícitas de cómo interviene la teoría de las funciones complejas. Más aún, a pesar de que Cauchy usó funciones complejas para evaluar integrales reales definidas, de la misma manera que lo habían hecho Euler y Laplace, este uso no manejaba funciones complejas como entidad básica. Tan tarde como en 1821, en su *Cours d'analyse (Curso de Análisis)* [14] dice que

$$\cos a+\sqrt{-1}\,\text{sen}\,a$$
$$\cos b+\sqrt{-1}\,\text{sen}\,b$$
$$\cos(a+b)+\sqrt{-1}\,\text{sen}\,(a+b)$$

«son tres expresiones simbólicas que no pueden ser interpretadas de acuerdo con las convenciones establecidas generalmente, y no representan algo real». El hecho que el producto de las anteriores expresiones primera y segunda iguala a la tercera, dice, no tiene sentido.

[12] *Œuvres*, (1), 1, 330.
[13] *Œuvres*, (1), 1, 338.
[14] *Œuvres*, (2), 3, 154.

Para dar sentido a esta ecuación hay que igualar la parte real y los coeficientes de $\sqrt{-1}$. «Cada ecuación imaginaria es únicamente la representación simbólica de dos ecuaciones entre cantidades reales». Si nosotros operamos sobre expresiones complejas de acuerdo con reglas establecidas para cantidades *reales*, obtenemos resultados exactos que son frecuentemente importantes.

En este libro trata de números complejos y variables complejas $u + \sqrt{-1}v$, donde u y v son funciones de *una* variable real, pero siempre en el sentido que las dos componentes reales son su contenido significante. Las funciones de valores complejos de una variable compleja no fueron consideradas.

En el año 1822, Cauchy dio algunos pasos hacia adelante. A partir de las relaciones (14) y (15) tenía

$$\int_{x_0}^{X} [V(x, Y) - V(x, y_0)]\, dx = \int_{y_0}^{Y} [S(X, y) - S(x_0, y)]\, dy \quad (16)$$

y

$$\int_{x_0}^{X} [S(x, Y) - S(x, y_0)]\, dx = -\int_{y_0}^{Y} [V(X, y) - V(x_0, y)]\, dy. \quad (17)$$

Ahora tuvo la idea de que podía combinar estas dos ecuaciones y entonces obtener resultados acerca de $F(z) = F(x + iy) = S + iV$. Así, multiplicando (16) por i y sumando las dos ecuaciones, obtiene

$$\int_{x_0}^{X} F(x + iY)\, dx - \int_{x_0}^{X} F(x + iy_0)\, dx$$

$$= \int_{y_0}^{Y} F(X + iy)i\, dy - \int_{y_0}^{Y} F(x_0 + iy)i\, dy,$$

las cuales, arreglando términos, dan

$$\int_{y_0}^{Y} F(x_0 + iy)i\, dy + \int_{x_0}^{X} F(x + iY\, dx$$

$$= \int_{x_0}^{X} F(x + iy_0)\, dx + \int_{y_0}^{Y} F(X + iy)i\, dy. \quad (18)$$

Este último resultado es el teorema de la integral de Cauchy para el caso sencillo de integración alrededor de la frontera de un rectángulo (Fig. 27.4). Se puede expresar este resultado como

$$\int_{ADC} F(z)\,dz = \int_{ABC} F(z)\,dz. \tag{19}$$

Esto es, la integral es independiente de la trayectoria.

Las ideas anteriores las da Cauchy en una nota de 1822, en su *Resumé des leçons sur le calcul infinitésimal (Resumen de las lecciones sobre el cálculo infinitesimal)* [15], y en una nota a pie de página del artículo de 1814 que fue publicado en 1827. A partir de estos últimos escritos podemos ver cómo fue Cauchy de las funciones reales a las complejas.

En 1825, Cauchy escribió otro ensayo. «Memóire sur les intégrales définies prises entre des limites imaginaires» («Memoria sobre las integrales definidas tomadas entre límites imaginarios»), pero no fue publicada hasta 1874 [16]. Este ensayo está considerado por muchos como el más importante, y uno de los más bellos en la historia de la ciencia, aunque por largo tiempo el propio Cauchy no apreció su valor.

En dicho texto considera nuevamente la cuestión de calcular integrales reales por el método de sustituir valores complejos por constantes y variables. Trata

$$\int_{x_0+iy_0}^{X+iY} f(z)\,dz, \tag{20}$$

donde $z = x + iy$, y define cuidadosamente la integral como el límite de la suma

$$\sum_{v=0}^{n-1} f(x_v + iy_v)[(x_{v+1} - x_v) + i(y_{v+1} - y_v)],$$

donde $x_0, x_1, ..., X$, e $y_0, y_1, ..., Y$ son los puntos de subdivisión a lo largo de la trayectoria a partir de (x_0, y_0) hasta (X, Y). Aquí $x + yi$

[15] *Œuvres*, (2), 4, 13-256.

[16] *Bull. des. Sci. Math.*, 7, 1874, 265-304 y 8, 1875, 148-159; este artículo no se encuentra en las *Œuvres* de Cauchy.

es claramente un punto del plano complejo y la integral se calcula sobre la trayectoria compleja. También muestra que si ponemos $x = \phi(t)$, $y = \psi(t)$, donde t es real, entonces el resultado es independiente de la elección de ϕ y ψ, lo que significa que es independiente de la trayectoria, siempre y cuando ninguna discontinuidad de $f(z)$ caiga entre dos trayectorias distintas. Este resultado generaliza los resultados para rectángulos.

Cauchy formula el teorema como sigue: si $f(x + iy)$ es finita y continua para $x_0 \leqslant x \leqslant X$ e $y_0 \leqslant y \leqslant Y$, entonces el valor de la integral (20) es independiente de la forma de las funciones $x = \phi(t)$ y $y = \psi(t)$. Su prueba del teorema usa el método del cálculo de variaciones. Considera como una trayectoria alternativa $\phi(t) + \varepsilon u(t)$, $\psi(t) + \varepsilon v(t)$, y muestra que la primera variación de la integral con respecto a ε es nula. Esta prueba no es satisfactoria. En ella Cauchy no sólo utiliza la existencia de la derivada de $f(z)$, sino también la continuidad de la derivada, a pesar de que no supone ninguno de los hechos en el enunciado de su teorema. La explicación es que Cauchy creía que una función continua siempre es diferenciable y que la derivada puede ser discontinua únicamente donde la propia función es discontinua. La creencia de Cauchy era razonable en tanto que al principio de su trabajo una función significaba para él, como para otros muchos en los siglos XVIII y XIX, una expresión analítica, y la derivada está dada inmediatamente por las reglas formales de diferenciación.

En el artículo de 1825 Cauchy es más claro aún acerca de una idea importante que ya había aparecido en el mismo artículo de 1814 y en una nota a pie de página a aquel ensayo. Considera lo que sucede cuando $f(z)$ es discontinua dentro o sobre la frontera del rectángulo (Fig. 27.4). Entonces el valor de la integral a lo largo de dos trayectorias diferentes puede ser diferente. Si en $z_1 = a + ib$, $f(z)$ es infinito, pero el límite

$$F = \lim_{z \to z_1} (z - z_1) f(z)$$

existe, esto es, si f tiene un polo simple en z_1, entonces la diferencia en las integrales es $\pm 2\pi\sqrt{-1}F$. Así, para la función $f(z) = 1/(1 + z^2)$, que es infinita en $z = \sqrt{-1}$, tal que $a = 0$ y $b = 1$,

$$F = \lim_{\substack{x \to 0 \\ y \to 1}} \frac{x + (y - 1)\sqrt{-1}}{[x + (y + 1)\sqrt{-1}][x + (y - 1)\sqrt{-1}]} = \frac{-\sqrt{-1}}{2}. \qquad (21)$$

La cantidad F es lo que Cauchy llamó el *résidu intégral* (residuo integral) en su *Exercices de mathématiques (Ejercicios de Matemáticas)* [17]. También, cuando una función tiene varios polos en la región limitada por las dos trayectorias de integración, Cauchy señala que uno debe tomar la suma de los residuos para obtener la diferencia de las integrales sobre las dos trayectorias. En esta sección, que trata en particular sobre residuos, sus dos trayectorias forman aún un rectángulo, pero toma uno muy grande y deja que los lados se hagan infinitos en longitud para que incluya todos los residuos.

En su *Exercices* [18], Cauchy señala que el residuo de $f(z)$ en z_1 es también el coeficiente del término $(z - a_1)^{-1}$ en el desarrollo de $f(z)$ en potencias de $z - a_1$. Mucho más tarde, en un artículo de 1841 [19], Cauchy proporcionó una nueva expresión para el residuo en el polo, a saber,

$$F(z_1) = E[f(z)]_{z_1} = \frac{1}{2\pi i}\int f(z)\,dz,$$

donde la integral es tomada a lo largo de un pequeño círculo que encierra $z = z_1$. El concepto y desarrollo de los residuos es una contribución importante de Cauchy. El uso inmediato de todos sus resultados obtenidos hasta entonces fue para evaluar integrales definidas.

Es evidente, a partir de lo que Cauchy añadió en las notas a pie de página a su ensayo de 1814 y por lo que escribió en su magistral ensayo de 1825, que debió haber pensado mucho y profundamente para percatarse de que las relaciones entre pares de funciones reales logran sus formas más sencillas cuando son introducidas cantidades complejas. Cuánto conocía del asunto a partir de los trabajos de Gauss y Poisson, no se sabe.

Durante los años de 1830 a 1838, cuando vivía en Turín y en Praga, sus publicaciones se hicieron dispersas. Se refiere a ellas y repite la mayor parte del trabajo en sus *Exercices d'analyse et de physique mathématique* (*Ejercicios de análisis y de física matemática.* 4 vols., 1840-1847). En un ensayo de 1831, publicado más tarde [20],

[17] Cuatro vols., 1826-30 = *Œuvres*, (2), 6-9.

[18] Vol. 1, 1826, 23-37 = *Œuvres*, (2), 6, 23-37.

[19] *Exercices d'analyse et de physique mathématique (Ejercicios de análisis y de física matemática)*, Vol. 2, 1841, 48-112 = *Œuvres*, (2), 12, 48-112.

[20] *Comp. Rend.*, 4, 1837, 216-218 = *Œuvres*, (1), 4, 38-42; véase también *Exer-*

obtuvo el siguiente teorema: la función $f(z)$ puede ser expandida de acuerdo con la fórmula de Maclaurin en una serie de potencias que es convergente para toda z cuyo valor absoluto es más pequeño que aquellos para los cuales la función o sus derivadas cesan de ser finitas o continuas. (Las únicas singularidades que Cauchy conocía son las que ahora nosotros llamamos polos.) Muestra que esta serie es menor término a término que una serie geométrica convergente cuya suma es

$$\frac{Z}{Z - z}\overline{f(z)},$$

donde Z es el primer valor para el cual $f(z)$ es discontinua y $\overline{f(z)}$ es el valor máximo de $|f(z)|$ para toda z cuyo valor absoluto iguala $|Z|$. De esta manera, Cauchy proporciona un criterio poderoso y fácil de aplicar para el desarrollo de una función en serie de Maclaurin que usa una serie de comparación ahora llamada serie mayorante.

En la prueba del teorema, primero muestra que

$$f(z) = \frac{1}{2\pi}\int_{\pi}^{\pi} \frac{\bar{z}f(\bar{z})}{\bar{z} - z}\, d\phi,$$

donde $\bar{z} = |Z|\, e^{i\varphi}$. Este resultado es prácticamente lo que ahora llamamos la fórmula integral de Cauchy. Más adelante expande la fracción $\bar{z}/(\bar{z} - z)$ en una serie geométrica en potencias de $z/\bar{z}$ y prueba el teorema.

En este teorema Cauchy también supone la existencia y continuidad de la derivada como siguiéndose necesariamente de la continuidad de la función misma. En el momento en que reprodujo este material en su *Exercices,* había intercambiado correspondencia con Liouville y Sturm y añadió al argumento del teorema anterior que la región de convergencia acaba en el valor de z para el que la función y su derivada cesan de ser finitos o continuos. Pero no estaba convencido que se deba añadir condiciones sobre la derivada, y en trabajos posteriores las eliminó.

En otro trabajo importante sobre la teoría de funciones complejas, «Sur les intégrales qui s'étendent à tous les points d'une courbe

---

cices d'analyse et de physique mathématique, Vol. 2, 1841, 48-112 = *Œuvres*, (2), 12, 48-112.

fermée» («Sobre las integrales que se extienden a todos los puntos de una curva cerrada»)[21] Cauchy relaciona la integral de una $f(z) = u + iv$ analítica alrededor de una curva encerrando un área [simplemente conexa] a una integral sobre el área. De esta manera si u y v son funciones de x e y,

$$\iint \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\, dy = \int u\, dy + \int v\, dx$$

$$\iint \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) dx\, dy = \int -u\, dx + \int v\, dy, \qquad (22)$$

donde las integrales dobles son sobre el área y las integrales sencillas sobre la frontera de la curva. Ahora, a la vista de las ecuaciones de Cauchy-Riemann (*cf.* [13]), los lados izquierdos son 0 y los dos lados derechos son las integrales que aparecen en

$$\int f(z)\, dz = \int (u + iv)(dx + i\, dy) = \int (u\, dx - v\, dy) + i \int (u\, dy + u\, dx).$$

De aquí que $\int f(z)\, dz = 0$, y Cauchy tiene una nueva prueba del teorema básico sobre la independencia de la trayectoria. Prueba el teorema para un rectángulo y en seguida lo generaliza para una curva cerrada que no se corta a sí misma. (El teorema también fue obtenido independientemente por Weierstrass en 1842.) No es seguro si Cauchy obtuvo esta formulación mucho más fértil de sus primeras ideas mediante la lectura del trabajo de Green en 1828 (cap. 28, sec. 4), pero existen indicaciones en esta dirección, ya que Cauchy extendió el resultado anterior a áreas sobre superficies curvadas.

En el artículo de 1846 mencionado con anterioridad y en otro del mismo año[22], Cauchy cambió su punto de vista sobre las funciones complejas, en contra de su trabajo de 1814, 1825 y 1826. En lugar de estar interesado en las integrales definidas y su cálculo, pasó a la teoría de funciones complejas en sí misma y a construir una base

[21] *Comp. Rend.*, 23, 1846, 251-255 = *Œuvres*, (1), 10, 70-74.

[22] «Sur les intégrales dans lesquelles la fonction sous le signe ∫ change brusquement de valeur» («Sobre las integrales en las que la función bajo el signo ∫ cambia bruscamente de valor»), *Comp. Rend.*, 23, 1846, 537 y 557-569 = *Œuvres*, (1), 10, 133-134 y 135-143.

para esta teoría. En este segundo artículo de 1846, proporciona un nuevo argumento sobre la $\int f(z)\, dz$ a lo largo de una curva arbitraria cerrada: si la curva encierra polos, entonces el valor de la integral es $2\pi i$ veces la suma de los residuos de la función en estos polos; esto es,

$$\int f(z)\, dz = 2\pi i E[f(z)], \tag{23}$$

donde $E[f(z)]$ es su notación para la suma de los residuos.

También tomó el tema de las integrales de funciones con valores múltiples [23]. En la primera parte del ensayo, donde trata integrales de funciones con un valor único, no dice mucho más de lo que ya había señalado Gauss en su carta a Bessel a propósito de $\int dx/x$ o $\int dx/(1+x^2)$. Claro, las integrales tienen valores múltiples y sus valores dependen de la trayectoria de integración. Pero Cauchy va más allá para considerar funciones con valores múltiples bajo el signo integral. Aquí afirma que si el integrando es una expresión para las raíces de una ecuación algebraica o trascendente, por ejemplo $\int w^3\, dz$ donde $x^3 = z$, y si uno integra sobre una trayectoria cerrada y regresa al punto inicial, entonces el integrando representa ahora otra raíz. En estos casos el valor de la integral sobre la trayectoria cerrada no es independiente del punto inicial, y el continuar alrededor de la trayectoria da diferentes valores de la integral. Pero si uno va alrededor de la trayectoria suficientes veces de tal forma que w regrese a su valor original, entonces los valores de la integral se repetirán y la integral es una función periódica de z. Los módulos de periodicidad *(índices de periodicité)* de la integral ya no son, como en el caso de funciones con un valor, representables por los residuos. Las ideas de Cauchy sobre las integrales de funciones multivaluadas eran aún demasiado vagas.

Durante veinticinco años, a partir de 1821, Cauchy desarrolló la teoría de funciones complejas por sí solo. En 1843, algunos compatriotas empezaron a tomar algún interés en su trabajo. Pierre-Alphonse Laurent (1813-1853), quien trabajaba solo, publicó un resul-

---

[23] «Considérations nouvelles sur les intégrales définies qui s'étendent à tous les points d'une courbe fermée» («Nuevas consideraciones sobre las integrales definidas que se extienden a todos los puntos de una curva cerrada»), *Comp. Rend.*, 23, 1846, 689-702 = *Œuvres*, (1), 10, 153-168.

tado importante obtenido en 1843. Mostró [24] que dada una función discontinua en un punto aislado, en lugar de una expansión de Taylor uno debe utilizar una expansión con potencias crecientes y decrecientes de la variable. Si la función y su derivada son univalentes y continuas en un anillo cuyo centro es un punto aislado a, entonces la integral de la función tomada sobre las dos fronteras circulares del anillo, pero en direcciones opuestas y debidamente expandidas, proporciona una expansión convergente dentro del propio anillo y con potencias crecientes y decrecientes de z. Esta expresión de Laurent es

$$f(z) = \sum_{-\infty}^{\infty} a_n(z - a)^n \tag{24}$$

y es una extensión del desarrollo de Taylor. El resultado le era conocido a Weierstrass en 1841, pero no lo publicó [25].

El tema de estas funciones multivaluadas fue retomado por Victor-Alexandre Puiseux. En 1850, Puiseux publicó un famoso ensayo [26] sobre funciones algebraicas complejas dadas por $f(u,z) = 0$, siendo f un polinomio en u y z. Primero hizo una clara distinción entre polos y puntos de ramificación que Cauchy apenas percibió e introdujo la noción de un punto singular esencial (un polo de orden infinito), sobre el que Weierstrass había llamado la atención independientemente. Tal punto es ejemplificado por $e^{1/z}$ en $z = 0$. A pesar de que Cauchy había considerado en el artículo de 1846 la variación de funciones multivaluadas simples a lo largo de trayectorias que encerraban puntos de ramificación, Puiseux clarificó también esta cuestión. Muestra que si u_1 es una solución de $f(u,z) = 0$ y z varía a lo largo de alguna trayectoria, el valor final de y_1 no depende de la trayectoria, con tal que la trayectoria no encierre algún punto en el cual u_1 es infinita o algún punto donde u_1 es igual a alguna otra solución (esto es, un punto de ramificación).

Puiseux también mostró que el desarrollo de una función de z alrededor de un punto de ramificación $z = a$ debe incluir potencias fraccionarias de $z - a$. Más adelante mejoró el teorema de Cauchy

[24] *Comp. Rend.*, 17, 1843, 348-349; el artículo completo fue publicado en el *Jour. de l'Ecole Poly.*, 23, 1863, 75-204.

[25] *Werke*, 1, 51-66.

[26] *Jour. de Math.*, 15, 1850, 365-480.

sobre la expansión de una función en serie de Maclaurin. Puiseux obtuvo una expansión para una solución u de $f(u,z) = 0$ no en potencias de z sino en potencias de $z - c$, y, por lo tanto, válida en un círculo con c como centro y sin contener ningún polo ni punto de ramificación. Después, Puiseux permite a c variar a lo largo de la trayectoria de manera que los círculos de convergencia coinciden en forma tal que el desarrollo dentro de un círculo puede ser extendido a otro. De esta manera, empezando con un valor de u en cualquier punto, uno puede seguir su variación a lo largo de cualquier trayectoria.

Mediante sus importantes investigaciones sobre funciones multivaluadas y sus puntos de ramificación en el plano complejo, y por su trabajo inicial sobre integrales de tales funciones, Puiseux llevó el trabajo inicial de Cauchy en teoría de funciones al final de lo que podría ser llamado la primera etapa. Las dificultades en la teoría de funciones multivaluadas y las integrales de tales funciones aún tenían que ser superadas. Cauchy escribió otros artículos sobre las integrales de funciones multivaluadas [27], en los que intentaba seguir el trabajo de Puiseux; y a pesar de que introdujo la noción de líneas de corte (lignes d'arrêt), aún estaba confundido acerca de la distinción entre polos y puntos de ramificación. Esta materia de las funciones algebraicas y sus integrales habría de continuarla Riemann (sec. 8).

En otros artículos de las *Comptes Rendus* de 1851 [28], Cauchy ofreció algunos argumentos más cuidadosos sobre las propiedades de las funciones complejas. En particular, afirmó que la continuidad de las derivadas, así como también la continuidad de la función compleja misma, es necesaria para la expansión en series de potencias. También señaló que la derivada en $z = a$ de u como una función de z es independiente de la dirección en el plano $x + iy$ cuando z se aproxima a a, y que u satisface $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$.

En estos artículos de 1851, Cauchy introdujo nuevos términos. Usó *monotypique* y también *monodrome* cuando la función era univalente para cada valor de z en algún dominio. Una función es *monogène* si para cada z solamente tiene una derivada (esto es, la derivada es independiente de la trayectoria). Una función monódroma y monógena que nunca es infinita es llamada *synectique*. Más tarde,

[27] *Comp. Rend.*, 32, 1851, 68-75 y 162-164 = *Œuvres*, (1), 11, 292-300 y 304-305.

[28] *Œuvres*, (1), 11.

Charles A. A. Briot (1817-1882) y Jean-Claude Bouquet (1819-1885) introdujeron «holomorfa» en lugar de *synectique* y «meromorfa» si la función poseía unicamente polos en el dominio.

5. La visión de Weierstrass de la teoría de funciones

Mientras que Cauchy desarrollaba la teoría de funciones sobre la base de las derivadas e integrales de funciones representadas por expresiones analíticas, Karl Weierstrass desarrolló una nueva visión. Nacido en Westphalia en 1815, Weierstrass ingresó en la Universidad de Bonn a estudiar leyes. Después de cuatro años, se cambió a las matemáticas en 1838, pero no completó su doctorado. En su lugar se aseguró una licencia del estado para llegar a ser profesor de *gymnasium* (instituto) y desde 1841 a 1854 enseñó a los jóvenes materias tales como gramática y gimnasia. Durante estos años no tuvo contacto alguno con el mundo matemático, a pesar de que trabajó duro en la investigación matemática. Los pocos resultados que publicó durante este período le aseguraron un puesto enseñando materias de índole técnica en el Instituto Industrial de Berlín. En ese mismo año ingresó en la Universidad de Berlín y llegó a profesor en 1864. Se mantuvo en esta posición hasta su muerte, en 1897.

Fue un hombre metódico y esmerado. Contrariamente a Abel, Jacobi y Riemann, no tuvo destellos de intuición. De hecho, desconfiaba de la intuición e intentó poner el razonamiento matemático sobre una base firme. Mientras que la teoría de Cauchy descansaba sobre unos fundamentos geométricos, Weierstrass intentó construir la teoría de los números reales; cuando esto estuvo hecho (Cap. 41, sec. 3), alrededor de 1841, construyó la teoría de funciones analíticas sobre la base de las series de potencias, técnica que aprendió de su maestro Christof Gudermann (1798-1852), y el proceso de prolongación analítica. Este trabajo fue llevado a cabo en los cuarenta, aunque no lo publicó en aquel entonces. Contribuyó en muchos otros temas en la teoría de funciones y trabajó en el problema de los *n* cuerpos en astronomía y en la teoría de la luz.

Es difícil fechar las creaciones de Weierstrass, ya que muchas no las publicó una vez logradas. Mucho de lo que había hecho fue conocido por el mundo matemático a través de sus clases en la Universidad de Berlín. Cuando publicó su *Werke (Obras)* en los 1890, no se preocupó por la prioridad, ya que muchos de sus resultados,

mientras tanto, habían sido editados por otros. El se interesaba más por presentar su método para desarrollar la teoría de funciones.

El uso de las series de potencias para representar funciones complejas dadas en forma analítica ya era conocida. Sin embargo, la tarea —dada una serie de potencias que define una función en un dominio restringido, derivar otra serie de potencias que define la misma función en otros dominios sobre la base de teoremas de series de potencias— fue atacada por Weierstrass. Una serie de potencias en $z - a$ que es convergente en un círculo C de radio r cerca de a representa una función que es analítica en cada valor de z en el círculo C. Escogiendo un punto b en el círculo y usando los valores de la función y sus derivadas dados por la serie original, es factible obtener una nueva serie de potencias en $z - b$ cuyo círculo de convergencia C' corta al primer círculo. En los puntos comunes a los dos círculos las dos series proporcionan el mismo valor de la función. Sin embargo, en los puntos de C' que son exteriores a C, los valores de la segunda serie son una continuación analítica de la función definida por la primera serie. Prolongando tanto como fuera posible a partir de C' sucesivamente a otros círculos, se obtiene la continuación analítica completa $f(z)$. La $f(z)$ completa es la colección de valores de los puntos z en todos los círculos. Cada serie se llama un elemento de la función.

Es posible que, durante la extensión del dominio de la función por la adjunción de más y más círculos de convergencia, uno de los nuevos círculos pueda cubrir parte de un círculo que no lo preceda inmediatamente en la cadena, y los valores de la función en esta parte común del nuevo círculo y uno anterior pueden no coincidir. Entonces la función es multivaluada.

Los puntos singulares (polos o puntos de ramificación), que pudieran surgir en este proceso, que necesariamente están sobre los límites de los círculos de convergencia de las series de potencias, son incluidos en la función por Weierstrass si el orden de un punto singular es finito, ya que en dicho punto es posible una expansión en potencias de $(z - z_0)^{1/n}$ teniendo únicamente un número finito de términos con exponentes negativos. Para obtener expansiones cerca de $z = \infty$, Weierstrass usa series en $1/z$. Si el elemento de función converge en el plano entero, Weierstrass lo llama una función entera, y si no es una función entera racional, esto es, no es un polinomio, entonces tiene una singularidad esencial en ∞ (e.g. *sen* z).

Weierstrass también proporcionó el primer ejemplo de una serie

de potencias cuyo círculo de convergencia es su frontera natural; esto es, el círculo es una curva de puntos singulares, y un ejemplo de una expresión analítica que puede representar diferentes funciones analíticas en diferentes partes del plano.

6. Funciones elípticas

Durante la primera mitad del siglo, paralelamente al desarrollo de los teoremas básicos de la teoría de funciones complejas, hubo un desarrollo especial concerniente a las funciones elípticas y más tarde a las funciones abelianas. No hay duda de que Gauss obtuvo un buen número de resultados importantes en la teoría de funciones elípticas, ya que muchos de éstos fueron encontrados después de su muerte en artículos que no había publicado. Sin embargo, los fundadores reconocidos de la teoría de funciones elípticas son Abel y Jacobi.

Niels Henrik Abel (1802-1829) fue hijo de un pastor protestante pobre. Como estudiante en Cristianía (Oslo, Noruega), tuvo la suerte de contar con Berndt Michael Holboe (1795-1850) como maestro. Este último reconoció el genio de Abel y predijo, cuando Abel tenía diecisiete años de edad, que se convertiría en el mejor matemático de todo el mundo. Después de estudiar en Cristianía y en Copenhage, Abel recibió una beca que le permitió viajar. En París fue presentado a Legendre, Laplace, Cauchy y Lacroix, pero lo ignoraron. Habiendo agotado sus fondos, viajó a Berlín, donde residió los años de 1825 a 1827 con Crelle. Regresó a Cristianía tan cansado que le fue necesario, según él mismo escribió, asirse de la puerta de una iglesia. Para ganar dinero impartió clases a estudiantes jóvenes. Empezó a recibir mayor atención a través de sus trabajos publicados y Crelle pensó que le podría asegurar una plaza de profesor en la Universidad de Berlín. Pero Abel contrajo la tuberculosis y murió en 1829.

Abel conoció la obra de Euler, Lagrange y Legendre sobre integrales elípticas y pudo haber obtenido de ahí sugerencias para el trabajo que emprendió a partir de comentarios hechos por Gauss, especialmente en su *Disquisitiones Arithmeticae*. El mismo empezó a escribir artículos en 1825. Presentó su ensayo principal sobre integrales a la Academia de Ciencias de París el 30 de octubre de 1826, para ser publicado en su revista. Este ensayo, «Mémoire sur une

propriété générale d'une classe très-etendue de fonctions trascendantes» (Memoria sobre una propiedad general de una clase muy amplia de funciones trascendentes), contiene el gran teorema de Abel (sec. 7). Fourier, que era secretario de la Academia por aquellos días, leyó la introducción del artículo y refirió el trabajo a Legendre y Cauchy para ser evaluado, siendo este último el responsable principal. El artículo era largo y difícil, ya que contenía muchas ideas nuevas. Cauchy lo dejó de lado para favorecer su propio trabajo. Legendre lo olvidó. Después de la muerte de Abel, cuando su fama ya estaba bien establecida, la Academia buscó el artículo; encontrándolo en 1841, y fue publicado [29]. Abel editó otros ensayos sobre la teoría de ecuaciones y funciones elípticas en la revista de Crelle y en los *Annales* de Gergonne. Aparecieron a partir de 1827. Debido a que el artículo principal de Abel de 1826 no fue publicado hasta 1841, otros autores, al conocer los teoremas más limitados que salieron a la luz entre estas fechas, obtuvieron de forma independiente muchos de los resultados de Abel de 1826.

El otro descubridor de las funciones elípticas fue Carl Gustav Jacob Jacobi (1804-1851). A diferencia de Abel, vivió una vida callada. Nacido en Potsdam de familia judía, estudió en la Universidad de Berlín y en 1827 llegó a ser profesor en Königsberg. En 1842 tuvo que dejar su puesto por su mala salud. Le fue otorgada una pensión del gobierno prusiano y se retiró a Berlín, donde murió en 1851. Su fama fue grande aún en vida, y sus alumnos divulgaron sus ideas por muchos lugares.

Jacobi enseñó funciones elípticas durante muchos años. Su estilo se convirtió en el modelo según el cual se construyó la propia teoría de funciones. También trabajó en determinantes funcionales (Jacobianos), ecuaciones diferenciales ordinarias y parciales, dinámica, mecánica celeste, dinámica de fluidos y funciones e integrales hiperelípticas. Con frecuencia se le asigna a Jacobi la etiqueta de matemático puro, pero, como la mayoría de sus contemporáneos y matemáticos que le precedieron, tomó muy en serio la investigación de la naturaleza.

Mientras que Abel estaba trabajando en funciones elípticas, Jacobi, quien también había leído el trabajo de Legendre sobre integrales elípticas, empezó a trabajar sobre las funciones correspondien-

---

[29] *Mém. des sav. étrangers*, 7, 1841, 176-264 = *Œuvres*, 145-211.

tes. Presentó un escrito a la *Astronomische Nachrichten* [30], pero sin demostraciones. Casi simultáneamente, Abel publicó por su parte sus «Recherches sur les fonctions elliptiques» (Investigaciones sobre las funciones elípticas) [31]. Ambos habían llegado a la idea clave de trabajar con las inversas de las funciones integrales elípticas, idea que Abel poseía desde 1823. Después Jacobi proporcionó demostraciones de los resultados que había publicado en 1827 en diversos artículos en la revista de Crelle, entre los años 1828 y 1830. De ahí en adelante, ambos publicaron sobre funciones elípticas, pero mientras que Abel murió en 1829, Jacobi vivió hasta 1851 y pudo publicar mucho más. En particular, el trabajo de Jacobi *Fundamenta Nova Theoriae Functionum Ellipticarum (Fundamentos de la Nuea Teoría de Funciones Elípticas)* de 1829 [32] se convirtió en un texto clave sobre funciones elípticas.

A través de cartas de Jacobi, Legendre se familiarizó con los trabajos de Jacobi y Abel. Escribió a Jacobi el 9 de febrero de 1828: «es una gran satisfacción para mí el ver a dos jóvenes matemáticos que han cultivado tan exitosamente una rama del análisis que por mucho tiempo ha sido mi campo favorito, pero que no se ha recibido como merece en mi propio país». Entonces Legendre publicó tres suplementos (1829 y 1832) a su *Traité des fonctions elliptiques* (*Tratado de Funciones Elípticas,* 2 vols., 1825-1826), en los que describió los logros del trabajo de Jacobi y Abel.

La integral elíptica general es

$$u = \int R(x, \sqrt{P(x)}), \tag{25}$$

donde $P(x)$ es un polinomio de tercero o cuarto grado con raíces diferentes y $R(x,y)$ es una función racional de x e y. Los esfuerzos para deducir algunos hechos generales acerca de la función u de x tenían que fracasar porque el propio significado de la integral estaba limitado para Euler y Legendre. Los coeficientes de $P(x)$ eran reales, y el rango de x real y, más aún, no contenía una raíz de $P(x) = 0$. Con un mayor conocimiento de la teoría de las funciones complejas, algún avance pudo haberse logrado en saber algo acerca de u como

[30] *Astron. Nach.*, 6, 1827, 33-38 = *Werke,* 1, 31-36.
[31] *Jour. für Math.*, 2, 1827, 101-181, y 3, 1828, 160-190 = *Œuvres,* 263-388.
[32] *Werke,* 1, 49-239.

función de x, pero este conocimiento no estaba a mano. Como resultó después, Abel y Jacobi tuvieron una idea mejor.

Para ser precisos, Legendre había introducido (cap. 19, sec. 4) las integrales elípticas $F(k, \phi)$, $E(k,\phi)$ y $\pi(n,k,\phi)$. Fue Abel quien, alrededor de 1826, observó que si, para considerar por ejemplo $F(k,\phi)$, uno estudiaba

$$u = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \int_0^\phi \frac{d\phi}{\sqrt{1-k^2 \operatorname{sen}^2 \phi}} \tag{26}$$

donde $x = sen\phi$, entonces se encontraban las mismas dificultades que cuando uno estudiaba

$$u = \int_0^x \frac{dx}{\sqrt{1-x^2}} = \operatorname{arc\,sen} x.$$

Las relaciones más hermosas vinieron de estudiar x como función de u. Entonces, Abel propuso estudiar x como función de u en el caso de las integrales elípticas. Ya que $x = sen\phi$, uno también puede estudiar ϕ como función de u.

Jacobi introdujo [33] la notación $\phi = am\ u$ para la función $|\phi|$ de u definida por (26). También introdujo

$$\cos \phi = \cos am\ u \ \text{ y } \ \Delta\phi = \Delta am\ u = \sqrt{1-k^2 \operatorname{sen}^2 \phi}.$$

Esta notación fue abreviada por Gudermann a

$$x = \operatorname{sen} \phi = \operatorname{sen} am\ u = sn\ u, \quad \cos \phi = \cos am\ u = cn\ u,$$
$$\Delta\phi = \Delta am\ u = dn\ u.$$

Tenemos inmediatamente que

$$sn^2\ u + cn^2\ u = 1, \quad dn^2\ u + k^2\ sn^2\ u = 1.$$

Si ϕ es cambiada en $-\phi$, entonces u cambia de signo. De aquí que

$$am(-u) = -am\ u, \quad sn(-u) = -sn\ u, \quad cn(-u) = cn\ u,$$
$$dn(-u) = dn\ u.$$

---

[33] *Fundamenta Nova*, 1829.

El papel de π en las funciones trigonométricas aquí es jugado por la cantidad K definida por

$$K=\int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-k^2 \operatorname{sen}^2 \phi}} = F\left(k, \frac{\pi}{2}\right).$$

Asociada a K está la cantidad trascendente K', que es la misma función de k' que K es de k, y donde k' está definida por $k^2 + k'^2 = 1$, $0 < k < 1$.

Lo que es importante (pero que no demostrado allí) acerca de K y K' es que

$$sn(u \pm 4K) = sn\, u, \quad cn(u + 4K) = cn\, u, \quad dn(u \pm 2K) = dn\, u.$$

De aquí que $4K$ es un período de las funciones elípticas $sn\, u$ y $cn\, u$, y $2K$ es un período de $dn\, u$.

Las funciones $sn\, u$, $cn\, u$ y $dn\, u$ están definidas de esta manera para x y u reales. Abel tenía lo que él consideraba como un elemento de cada función, ya que cada uno está definido únicamente para valores reales. Su siguiente idea fue definir las funciones elípticas en su totalidad mediante la introducción de valores complejos de u. En cuanto al conocimiento de funciones complejas, en su visita a París, Abel conoció el trabajo de Cauchy. De hecho ya había estudiado el teorema del binomio para valores complejos de la variable y del exponente. La extensión primero a valores imaginarios puros fue lograda mediante lo que se llama la transformación imaginaria de Jacobi. Abel introdujo

$$\operatorname{sen} \theta = i \tan \phi, \qquad \cos \theta = \frac{1}{\cos \phi}, \qquad \Delta(\theta, k) = \frac{\Delta(\phi, k')}{\cos \phi},$$

donde $\theta = am\, i\, u$, de tal modo que

$$sn(iu, k)=i\,\frac{sn(u, k')}{cn(u, k')}, \; cn(iu, k)=\frac{1}{cn(u, k')}, \; dn(iu, k)=\frac{dn(u, k')}{cn(u, k')}.$$

Además de permitir a sus variables tomar valores imaginarios puros, Abel también desarrolló lo que es denominado teorema de adición para funciones elípticas. En el caso donde

$$u = A(x) = \int_0^x \frac{1}{\sqrt{1-x^2}}\,dx$$

sabemos que la integral es una función multivaluada $A(x) = arc\ sen\ x$ y es cierto que

$$A(x_1) + A(x_2) = A(x_1y_2 + x_2y_1), \tag{27}$$

donde las y_1 y y_2 son los valores correspondientes del coseno; i.e. $y_1 = \sqrt{1 - x_1{}^2}$. Pero en este caso se obtiene una gran simplificación al introducir la función inversa $x = sen\ u$, que es univalente, y en lugar de (27) tenemos el teorema familiar de la suma para la función seno. Ahora, en el caso de

$$u = E(x) = \int_0^x \frac{dx}{\sqrt{R(x)}},$$

donde $y^2 = R(x)$ es un polinomio de cuarto grado, Euler había obtenido el teorema de la adición (cap. 19, sec. 4)

$$E(x_1) + E(x_2) = E(x_3),$$

donde x_3 es una función racional conocida de x_1, x_2, y_1, y_2 e $y = \sqrt{R(x)}$. Abel pensó que para la función inversa $x = \phi(u)$ debía existir un teorema sencillo de suma y éste fue el caso. Dicho resultado aparece también en su ensayo de 1827. Se deduce entonces que para u y v reales

$$sn(u+v) = \frac{sn\ u\ cn\ v\ dn\ v + sn\ v\ cn\ u\ dn\ u}{1 - k^2 sn^2\ u\ sn^2\ v} \tag{28}$$

con fórmulas análogas para $cn(u + v)$ y $dn(u + v)$. Estos son los teoremas aditivos para funciones elípticas y los análogos de los teoremas aditivos para integrales elípticas.

Habiendo definido las funciones elípticas para valores reales e imaginarios de los argumentos, Abel, con los teoremas aditivos, fue capaz de extender las definiciones a valores complejos. Porque, si $z = u + iv$, $sn\ z = sn(u + iv)$ tiene significado en función del teorema aditivo, en términos de las sn, cn y dn de u y de iv separadamente.

También se sigue que

$$\begin{aligned} sn(iu + 2iK', k) &= sn(iu, k) \\ cn(iu + 4iK', k) &= cn(iu, k) \\ dn(iu + 4iK', k) &= dn(iu, k). \end{aligned} \tag{29}$$

De esta manera los períodos (que no son únicos) de $sn\ z$ son $4K$ y $2iK'$; los de cn z son $4K$ y $2K + 2iK'$; y los de $dn\ z$ son $2K$ y $4iK'$. Lo importante acerca de los períodos es que hay dos períodos (cuya razón no es real), de tal manera que las funciones son doblemente periódicas. Este fue uno de los resultados más importantes de Abel. Las funciones son univalentes. Necesitan, por tanto, ser estudiadas únicamente en un paralelogramo (Fig. 27.5) del plano complejo, ya que éstas repiten su comportamiento en todo paralelogramo congruente. Las funciones elípticas, además de ser univalentes y doblemente periódicas, tienen una única singularidad esencial en ∞. De hecho estas propiedades pueden utilizarse para definir las funciones elípticas, que tienen polos dentro de cada paralelogramo periódico.

A pesar de que Abel había arrebatado a Legendre lo que podía haber sido lo mejor de su vida de trabajo al introducir la idea de la inversión de las integrales elípticas, que Legendre pasó por alto y demostró ser la clave para explorarlas, Legendre elogió a Abel diciendo: «Qué cabeza tiene este noruego». Charles Hermite comentó que Abel había dejado ideas sobre las que podían trabajar los matemáticos por 150 años.

Muchos de los resultados de Abel fueron obtenidos aparte por Jacobi, cuya primera publicación en este campo —mencionada anteriormente— apareció en 1827. Jacobi era consciente del hecho de que el método básico que había utilizado en su *Fundamenta Nova*

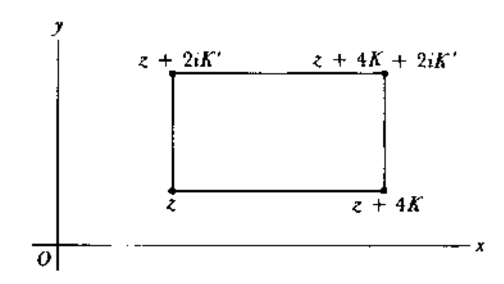


FIGURA 27.5

era insatisfactorio, y en cierto grado, en ese libro y en conferencias subsecuentes, usó un punto de partida diferente. Sus conferencias nunca fueron publicadas por completo, pero su contenido esencial es bastante conocido a través de cartas y notas dirigidas a sus alumnos. En su nueva manera, construyó su teoría de las funciones elípticas sobre la base de funciones auxiliares llamadas funciones theta, que están ilustradas por la

$$\theta(z) = \sum_{n=-\infty}^{\infty} e^{-n^2t+2ntz}, \tag{30}$$

donde x y t son complejos y $Re(t) > 0$. La serie converge absolutamente y uniformemente en cualquier region acotada del plano z. Jacobi introdujo cuatro funciones theta y más adelante expresó *sn u*, *cn u* y *dn u* en términos de estas funciones. Las funciones theta son los elementos más simples a partir de los cuales pueden ser construidas las funciones elípticas. Asimismo obtuvo varias expresiones para las funciones theta en forma de series y productos infinitos. Su trabajo, basado en ideas de Abel, llevó a Jacobi a obtener relaciones entre las funciones theta y la teoría de números. Esta conexión fue seguida posteriormente por Hermite, Kronecker y otros. El estudio de relaciones entre muy diversas formas de funciones theta constituyó una de las principales actividades de los matemáticos del siglo XIX; se trató de una de las muchas modas que surgieron con regularidad dentro de las matemáticas.

En un artículo importante en 1835 [34], Jacobi demostró que una función univaluada de una variable que para cualquier valor finito del argumento tiene el carácter de una función racional (lo cual significa que es una función meromorfa) no puede tener más de dos períodos y la razón de los períodos es necesariamente un número que no es real. Este descubrimiento abrió una nueva dirección de trabajo, a saber, el problema de encontrar todas las funciones doblemente periódicas. En 1844 [35] Liouville, en una comunicación a la Academia Francesa de Ciencias, mostró cómo desarrollar una teoría completa de funciones elípticas doblemente periódicas a partir del teorema de Jacobi. Esta teoría fue una contribución importante a las funciones elípticas. En la periodicidad doble Liouville había descu-

[34] *Journ. für Math.*, 13, 1835, 55-78 = *Werke*, 2, 23-50.

[35] *Comp. Rend.*, 19, 1844, 1261-1263, y 32, 1851, 450-452.

bierto una propiedad esencial de las funciones elípticas y un punto de vista que unificaba la teoría, pese a que las funciones doblemente periódicas son una clase más general que aquellas designadas por Jacobi como elípticas. Sin embargo, todas las propiedades fundamentales de las funciones elípticas se mantienen para las funciones doblemente periódicas.

Weiestrass retomó las funciones elípticas alrededor de 1860. Supo del trabajo de Jacobi mediante el de Gudermann, y del trabajo de Abel a través de sus artículos, los cuales le impresionaron tanto que constantemente recomienda la lectura de Abel a sus alumnos. Para su certificado de maestro tomó un problema que le había dado Gudermann: representar funciones elípticas como cocientes de series de potencias. Lo hizo. Como profesor, en sus clases retrabajó constantemente su teoría de las funciones elípticas.

Legendre había reducido las integrales elípticas a tres formas canónicas que incluían la raíz cuadrada de un polinomio de cuarto grado. Weierstrass llegó a tres formas diferentes con la raíz cuadrada de un polinomio de tercer grado [36], a saber

$$\int \frac{dx}{\sqrt{4x^3-g_2x-g_3}}, \int \frac{x\,dx}{\sqrt{4x^3-g_2x-g_3}}, \int \frac{dx}{(x-a)\sqrt{4x^3-g_2x-g_3}},$$

e introdujo, como función elíptica fundamental, aquella que resulta de «invertir» la primera integral. Esto es, si

$$u = \int_0^x \frac{dx}{\sqrt{4x^3-g_2x-g_3}},$$

entonces la función elíptica x de u es la de Weierstrass:

$$x = \wp(u) = \wp(u|g_2, g_3).$$

Para que $\wp(u)$ no degenere en una función exponencial o trigonométrica, es necesario que el discrimante $g^3{}_2 - 27\, g^2{}_3 \neq 0$, o, en otras palabras, que las tres raíces de la cúbica en x deben ser diferentes.

[36] *Sitzungsber. Akad. Wiss. zu Berlin,* 1882, 443-451 = *Werke,* 2, 245-255; véase también *Werke,* 5.

La $\wp(u)$ doblemente periódica de Weierstrass hace el papel de *sn u* en la teoría de Jacobi y es la más sencilla de las funciones doblemente periódicas. Demostró que toda función elíptica puede ser expresada simplemente en términos de $\wp(u)$ y la derivada de esta función. La «trigonometría» de las funciones elípticas es más sencilla en la formulación de Weierstrass, pero las funciones de Jacobi y las integrales de Legendre son mejores para el trabajo numérico.

De hecho, Weierstrass empezó con un elemento de su $\varphi(u)$, a saber,

$$\wp(u) = \frac{1}{u^2} + \frac{g_2}{4 \cdot 5} u^2 + \frac{g_3}{4 \cdot 7} u^4 + \ldots, \qquad (g_2, g_3 \text{ complejos}),$$

que obtuvo al resolver la ecuación diferencial para dx/du dada por la integral anterior, y usando el teorema de adición para $\wp(u)$ de manera similar a la de Abel para obtener la función completa. El trabajo de Weierstrass completó, remodeló y llenó de elegancia la teoría de las funciones elípticas.

A pesar de que no debemos entrar en detalles específicos, no podemos dejar las funciones elípticas sin mencionar el trabjo de Charles Hermite (1822-1901), quien era profesor en la Sorbonne y en la Ecole Polytechnique. A partir de sus días de estudiante, Hermite se ocupó siempre de las funciones elípticas. En 1892 escribió: «No puedo abandonar las funciones elípticas; donde la cabra está sujeta, ahí debe pastar». Obtuvo resultados básicos en la propia teoría y estudió la relación con la teoría de números. Aplicó las funciones elípticas a la solución de la ecuación polinómica de quinto grado y trató problemas de mecánica usando estas funciones. También es conocido por su demostración de la trascendencia de *e* y su introducción de los polinomios de Hermite.

7. Las integrales hiperelípticas y el teorema de Abel

El éxito obtenido en el estudio de las integrales elípticas (25) y las funciones correspondientes motivaron a los matemáticos a atacar un caso aún más difícil, las integrales hiperelípticas.

Las integrales hiperelípticas son de la forma

$$\int R(x, y)\, dx, \tag{31}$$

donde $R(x,y)$ es una función racional de x e y, $y^2 = P(x)$, y el grado de $P(x)$ es al menos cinco. Cuando $P(x)$ es de grado cinco o seis, se llamó ultraelípticas a las integrales a mediados del siglo XIX. Para resaltar los valores complejos es común escribir

$$\int R(u, z)\, dz; \tag{32}$$

y $P(z)$ es escrito a menudo como

$$u^2 \equiv P(z) = A(z - e_1) \dots (z - e_n). \tag{33}$$

Por supuesto, u es una función multívoca de z.

Entre las integrales de la forma (32) hay algunas que son finitas en todas partes. Las básicas son

$$u_1 = \int \frac{dz}{u}, u_2 = \int \frac{z\, dz}{u}, \dots, u_p = \int \frac{z^{p-1}\, dz}{u}, \tag{34}$$

donde u está dada por (33) y $p = (n - 2)/2$ o $(n - 1)/2$, según que n sea par o impar. Para $n = 6$ (y por lo mismo $p = 2$), hay dos integrales. Las integrales generales (32) tienen a lo más polos y puntos logarítmicamente singulares, esto es, puntos singulares que se comportan como $\log z$ en $z = 0$. Las integrales de la primera clase, esto es, aquellas que son finitas en todas partes y que no tienen puntos singulares, pueden ser expresadas siempre en términos de las p integrales (34), que son linealmente independientes.

Para el caso $n = 6$ (y $p = 2$) las integrales de la segunda clase están ejemplificadas por

$$\int \frac{z^2\, dz}{\sqrt{P(z)}}, \int \frac{z^3\, dz}{\sqrt{P(z)}}, \tag{35}$$

donde $Pz)$ es un polinomio de sexto grado. Las integrales de primera y segunda clase para $n = 6$ tienen cuatro períodos cada una.

Las integrales hiperelípticas son funciones del límite superior z si el límite inferior es fijo. Supongamos que denotamos una de estas funciones por w. Entonces, como en el caso de las integrales elípticas, uno podría formularse la pregunta de cuál es la función inversa z de w. Este problema lo atacó Abel, pero no lo resolvió. Más

adelante, lo atacó Jacobi [37]. Consideremos con Jacobi el caso particular de las integrales hiperelípticas

$$w = \int_0^z \frac{dz}{\sqrt{P(z)}}, \quad y \quad w = \int_0^z \frac{z\,dz}{\sqrt{P(z)}}, \tag{36}$$

donde $P(z)$ es un polinomio de quinto o sexto grado. Aquí la determinación de z como una función univaluada de w se mostró sin esperanza. De hecho, Jacobi demostró que la mera inversión de tales integrales con $P(z)$ de grado cinco no conducía a funciones monógenas. Las funciones inversas le parecían irrazonables a Jacobi, ya que en cada caso z como función de w tiene infinitos valores; tales funciones no fueron bien entendidas entonces.

Jacobi decidió considerar combinaciones de tales integrales. Guiado por el teorema de Abel (véase más adelante), cuyo enunciado conocía al menos por haber sido publicado por este tiempo, Jacobi hizo lo siguiente. Consideremos las ecuaciones

$$\int_0^{z_1} \frac{dz}{\sqrt{P(z)}} + \int_0^{z_2} \frac{dz}{\sqrt{P(z)}} = w_1 \tag{37}$$

$$\int_0^{z_1} \frac{z\,dz}{\sqrt{P(z)}} + \int_0^{z_2} \frac{z\,dz}{\sqrt{P(z)}} = w_2. \tag{38}$$

Jacobi tuvo éxito en demostrar que las funciones simétricas $z_1 + z_2$ y $z_1 z_2$ son, cada una, funciones univaluadas de w_1 y w_2 con un sistema de *cuatro* períodos. Las funciones z_1 y z_2 de las dos variables w_1 y w_2 pueden ser obtenidas entonces. También proporcionó un teorema de adición para estas funciones. Jacobi dejó muchos puntos incompletos. «Para el rigor gaussiano», dijo, «no tenemos tiempo».

El estudio de una generalización de las integrales elípticas e hiperelípticas fue iniciado por Galois, pero los pasos iniciales más significativos fueron dados por Abel en su artículo de 1826. Consideró (32), esto es,

$$\int R(u, z)\,dz, \tag{39}$$

[37] *Journ. für Math.*, 9, 1832, 394-403 = *Werke*, 2, 7-16, y *Jour. für Math.*, 13, 1835, 55-78 = *Werke*, 2, 325-350 y 516-521.

pero en lugar de (33), donde u y z están meramente ligadas por un polinomio como en $u^2 = P(z)$, Abel consideró una ecuación algebraica general en z y u,

$$f(u, z) = 0. \tag{40}$$

Las ecuaciones (39) y (40) definen lo que se llama una *integral abeliana,* que para aquel entonces incluía como casos especiales las integrales elípticas e hiperelípticas.

A pesar de que Abel no llevó muy lejos el estudio de las integrales abelianas, demostró un teorema clave en la materia. El teorema básico de Abel es una generalización muy amplia del teorema de adición para funciones elípticas (cap. 19, sec. 4). El teorema y su demostración están en el artículo de París de 1826 y su enunciado en el *Diario de Crelle* de 1829 [38]. Consideremos la integral

$$\int R(x, y)\, dx, \tag{41}$$

donde x e y están relacionadas mediante $f(x,y) = 0$, siendo f un polinomio en x e y. Abel escribe como si x e y fueran variables reales, aunque ocasionalmente aparecen números complejos. Hablando libremente, el teorema de Abel es éste: una suma de integrales de la forma (41) puede ser expresada en términos de p integrales de ese tipo más términos algebraicos y logarítmicos. Más aún, el número p depende unicamente de la ecuación $f(x,y) = 0$ y de hecho es el género de la ecuación.

Para obtener un enunciado mucho más preciso, sea y la función algebraica de x definida por la ecuación

$$f(x, y) = y^n + A_1 y^{n-1} + \ldots + A_n = 0, \tag{42}$$

donde las A_i son polinomios en x y el polinomio (42) es irreducible a factores de la misma forma. Sea $R(x,y)$ cualquier función racional de x e y. Entonces la suma de cualquier número m de integrales similares

$$\int^{(x_1,y_1)} R(x, y)\, dx + \ldots + \int^{(x_m,y_m)} R(x, y)\, dx \tag{43}$$

[38] *Jour. für Math.*, 4, 212-215 = *Œuvres*, 515-517.

con límites inferiores fijos (pero arbitrarios) es expresable mediante funciones racionales de x_1, $y_1,...,x_m$ e y_m y logaritmos de tales funciones racionales, con la adición de la suma de un cierto número p de integrales.

$$\int^{(z_1,s_1)} R(x,y)\,dx, \ldots, \int^{(z_p,s_p)} R(x,y)\,dx, \tag{44}$$

donde $z_1,...,z_p$ son valores de x determinables a partir de x_1, $y_1,...,x_m$, e y_m como las raíces de una ecuación algebraica cuyos coeficientes son funciones racionales de x_1, $y_1,...,x_m$ e y_m y $s_1,...,$ y s_p son los valores correspondientes de y y determinados por (42) con cualquier s_i determinable como función racional de las z_i y las x_1, $y_1,...,$ x_m e y_m. Las relaciones que así determinan $(z_1, x_1),..., (z_p, x_p)$ en términos de $(x_1, y_1),..., (x_m, y_m)$ debe suponerse que se cumplen en todos los pasos de la integración; en particular, estas relaciones determinan los límites inferiores de las últimas p integrales en términos de los límites inferiores de las primeras m integrales. El número p no depende de m ni de la forma de la función racional $R(x,y)$ ni de los valores x_1, $y_1,...,$ x_m y y_m, pero sí depende de la ecuación fundamental (42) que relaciona la y con la x.

En el caso de estas integrales hiperelípticas donde $f = y^2 - P(x)$ y $P(x)$ es un polinomio de grado sexto y donde p, que es $(n - 2)/2$, es 2, la parte principal del teorema de Abel dice que

$$\int_0^{x_1} R(x,y)\,dx + \ldots + \int_0^{x_m} R(x,y)\,dx$$

$$= \int_0^A R(x,y)dx + \int_0^B R(x,y)dx + R_1(x_1,y_1,\ldots,x_m,y_m,A,y(A),B,y(B))$$

$$+ \sum \text{const.} \log R_2(x_1, y_1, \ldots, x_m, y_m, A, y(A), B, y(B)), \tag{45}$$

donde R_1 y R_2 son funciones racionales de sus variables.

Abel, de hecho, calculó el número p para unos pocos casos del general $f(x,y) = 0$. A pesar de que no vio la verdadera importancia de este resultado sí reconoció la noción de género antes que Riemann y fundó las integrales abelianas. Su ensayo fue muy difícil de entender, en parte porque intentó demostrar lo que hoy en día lla-

maríamos un teorema de existencia mediante el cálculo del resultado. Demostraciones posteriores simplificaron considerablemente la de Abel (véase también Cap. 39, sec. 4). Abel no consideró el problema de la inversión. Todo el trabajo sobre la inversa de las integrales hiperelípticas y abelianas, hasta la llegada de Riemann a la escena, fue impedido por los métodos, tan limitados, de manejar las funciones multivaluadas.

8. Riemann y las funciones multívocas

Alrededor de 1850, un período de éxitos en la teoría de funciones llegó a su fin. Métodos rigurosos, tales como los que proporcionó Weierstrass, aclararon la dirección de los resultados, y las demostraciones incuestionables de existencia denotan en cualquier disciplina matemática una importante, pero última, etapa de su desarrollo. El desarrollo posterior debe estar precedido por un período de descubrimientos libres, numerosos, inconexos, por lo regular hechos por accidente y, tal vez, creaciones desordenadas. El teorema de Abel fue uno de esos pasos. A Riemann se debe un nuevo período de descubrimiento en la teoría de funciones algebraicas, sus integrales y sus funciones inversas. De hecho, Riemann ofreció una teoría mucho más amplia, a saber, el tratamiento de las funciones multivaluadas, hasta entonces únicamente tocadas por Cauchy y Puiseux, y a partir de entonces allanó el camino para un gran número de diferentes avances.

George Friedrich Bernhard Riemann (1826-1866) fue alumno de Gauss y de Wilhelm Weber. Llegó a Gottingen en 1846 para estudiar teología, pero rápidamente se cambió a matemáticas. Su tesis doctoral de 1851, escrita bajo la dirección de Gauss, y titulada «Grundlagen fur eine allgemeine Theorie des Functionen einer veranderlichen complexen Grosse» («Fundamentos de una teoría general de funciones de una variable compleja»)[39] es un ensayo básico de la teoría de funciones complejas. Tres años más tarde se convirtió en un *Privatdozent* en Gottingen, esto es, gozaba del privilegio de dar clases a estudiantes y cobrar una cuota. Para calificarse como *Privatdozent* tuvo que escribir su *Habilitationsschrift:* «Uber die Darstellbarkeit einer Function durch eine trigonometrische Reihe» («So-

---

[39] *Werke*, 3-43.

bre la representación de una función mediante una serie trigonométrica»), y dio una conferencia cualificadora, la *Habilitationsvortrag*, «Uber die Hypothesen welche der Geometrie zu Grunde liegen» («Sobre las hipótesis en las que está fundada la geometría»). Estas fueron seguidas por una serie de famosos artículos. Riemann fue el sucesor de Dirichlet como profesor de matemáticas en Gottingen en 1859. Murió de tuberculosis.

Riemann es frecuentemente descrito como un matemático puro, pero esto está muy lejos de ser cierto. A pesar de que hizo numerosas contribuciones a las matemáticas propiamente dichas, estaba profundamente interesado en la física y las relaciones de las matemáticas con el mundo físico. Escribió ensayos sobre el calor, la luz, la teoría de gases, el magnetismo, la dinámica de fluidos y la acústica. Intentó unificar la gravitación con la luz e investigar el mecanismo del oído humano. Su trabajo sobre los fundamentos de la geometría buscó asegurar lo que es absolutamente seguro acerca de nuestro conocimiento del mundo del espacio físico (cap. 37). El mismo Riemann asegura que su trabajo sobre las leyes físicas fue su interés primordial. Como matemático, usó libremente sus intuiciones geométricas y argumentos físicos. Parece posible, sobre la base de las pruebas dadas por Felix Klein, que las ideas de Riemann sobre funciones complejas le fueron sugeridas por sus estudios sobre el flujo de corrientes eléctricas a lo largo de un plano. La ecuación del potencial es central en esa materia y lo fue también en el acercamiento de Riemann a las funciones complejas.

La idea clave en la visión de Riemann de las funciones multivaluadas es la noción de superficie de Riemann. La función $w^2 = z$ es multivaluada y, de hecho, existen dos valores de w para cada valor de z. Para trabajar con esta función y mantener los dos conjuntos de valores $\sqrt{z}$ y $-\sqrt{z}$ aparte, esto es, para separar las ramas, Riemann introdujo un plano de valores de z para cada rama. Incidentalmente, también introdujo un punto sobre cada plano correspondiente a $z = \infty$. Los dos planos se consideran como si uno cayera sobre el otro y unidos, lo primero de todo, en aquellos valores de z donde las ramas dan los mismos valores de w. Así para $w^2 = z$ los dos planos, u hojas como son llamados, están unidos en $z = 0$ y $z = \infty$.

Ahora, $w = +\sqrt{z}$ está representada por los valores de z únicamente en la hoja superior y $w = -\sqrt{z}$ mediante los valores de z de la hoja inferior. Mientras que uno considere los valores de z de la

hoja superior, se entiende que es necesario calcular $w^1 = +\sqrt{z}$. Sin embargo, cuando z se mueve en un círculo alrededor del origen sobre esa hoja (fig. 27.6) de tal forma que el θ en $z = \varrho(\cos\theta + i \text{ sen } \theta)$ varía de 0 a 2 π, $\sqrt{z}$ cubre una mitad del círculo en el plano complejo en el cual los valores de w son aplicados. Ahora, dejemos que z se mueva en la segunda hoja cruzando, digamos, el eje x positivo. Así cuando z se mueve en la segunda hoja, tomamos como w los valores dados por $w_2 = -\sqrt{z}$. Cuando z describe otro circuito en torno al origen, del modo que 0 va de 2π a 4π sobre la segunda hoja, obtenemos el rango de los valores de $w_2 = -\sqrt{z}$ para esta trayectoria, y el ángulo polar de estos valores de w va de π a 2π. Cuando z cruza de nuevo el eje x positivo, lo consideramos como viajando sobre la primera hoja. Así, mediante dos vueltas de valores de z alrededor del origen, uno en cada hoja, obtenemos los valores de w de la función $w^2 = z$. Más aún, y esto es esencial, w se convierte en una función univaluada de z si z se mueve sobre una superficie de Riemann, la cual es el agregado de las dos hojas.

Para distinguir las trayectorias sobre una hoja de aquellas sobre la otra, acordamos en el caso de $w^2 = z$ considerar el eje x como una rama cortada. Esto une los puntos $z = 0$ y $z = \infty$. Esto significa que siempre que z cruce esta cortadura, esa rama de w debe ser tomada como perteneciente a la hoja por la que pasa z. La rama cortada debe no ser necesariamente el eje x pero debe, en el caso presente, unir 0 e ∞. Se dice que 0 e ∞ son puntos de ramificación porque las ramas de $w^2 = z$ se intercambian cuando z describe una curva cerrada alrededor de cada una de ellos.

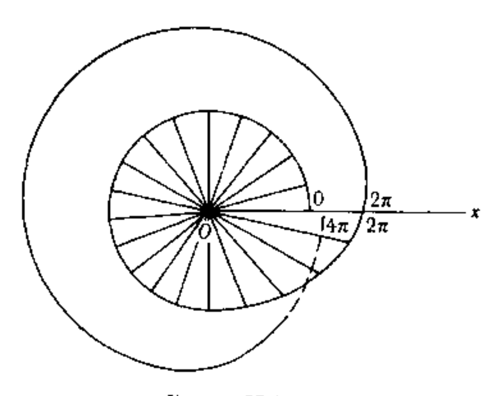


FIGURA 27.6

La función $w^2 = z$ y, consecuentemente, su superficie de Riemann asociada, son especialmente sencillas. Consideremos la función $w^2 = z^3 - z$. Esta función también tiene dos ramas que se hacen iguales a $z = 0$, $z = 1$, $z = -1$ y $z = \infty$. A pesar de que no presentaremos el argumento completo, los cuatro puntos son puntos de ramificación, ya que si z realiza un circuito alrededor de cualquiera de ellos, el valor de w cambia a partir de ese de una rama a otra. Las cortaduras de las ramas pueden ser tomadas como los segmentos de 0 a 1, 0 a -1, 1 a ∞ y -1 a ∞. Cuando z cruza cualquiera de estas cortaduras, el valor de w cambia de aquellos que toma en una rama a los de la segunda.

Para funciones multivaladas mucho más complicadas, la superficie de Riemann es más complicada. Una función con n valores requiere de una superficie de Riemann con n hojas. Puede haber muchos puntos de ramificación, y se deben introducir cortaduras de rama uniendo cada dos. Es más, las hojas que se unen en algún punto de ramificación no son necesariamente las mismas que las que se unen en otro. Si k hojas coinciden en algún punto, se dice que el orden del punto es $k - 1$. Sin embargo, dos hojas de una superficie de Riemann pueden tocarse en un punto, pero las ramas de la función es posible que permanezcan sin cambiar cuando z gira completamente en torno al punto. Entonces el punto no es un punto de ramificación.

No es posible representar fielmente las superficies de Riemann en un espacio tridimensional. Por ejemplo, las dos hojas para $w^2 = z$ deben intersecarse a lo largo del eje x positivo si es representada en tres dimensiones, de modo que uno debe ser cortado a lo largo del eje x positivo, mientras que las matemáticas requieren un paso suave de la primera hoja a la segunda y, entonces, después de una vuelta alrededor de $z = 0$, se regresa de nuevo a la primera hoja.

Las superficies de Riemann no sólo son una manera de representar funciones mutivaluadas, sino que, en efecto, es la forma de hacer tales funciones univaluadas sobre la superficie, como opuesta al plano de las z. Con ello, teoremas acerca de funciones univaluadas pueden ser extendidos a funciones multivaluadas. Por ejemplo, el teorema de Cauchy acerca de integrales de funciones *univaluadas* estando 0 sobre una curva limitando un dominio (en el que la función es analítica) fue extendido por Riemann a funciones multivaluadas. El dominio de analiticidad debe ser simplemente conexo (contractible en un punto) sobre la *superficie.*

Riemann pensaba su superficie como una duplicación de n hojas de plano, completada cada réplica por un punto en el infinito. Sin embargo, es difícil seguir todos los argumentos para tal superficie visualizándolos en términos de los n planos interconectados. De aquí que los matemáticos, desde los tiempos de Riemann, hayan sugerido modelos equivalentes que son más fáciles de contemplar. Es sabido que un plano puede ser transformado en una esfera mediante una proyección estereográfica (Cap. 7, sec. 5). Por eso podemos construir un modelo de la superficie de Riemann considerando n esferas concéntricas de aproximadamente el mismo radio. La sucesión de esferas es la misma que la de las hojas del plano. Los puntos de ramificación de estos planos y las cortaduras de las ramas son de la misma manera transformadas a esferas, de forma tal que estas esferas se envuelven una a otra a lo largo de cortaduras de ramas. Ahora pensamos el conjunto de esferas como el dominio de z, y la función multivaluada w de z es univalente en este conjunto de hojas esféricas.

En nuestra exposición de las ideas de Riemann, hemos empezado con una función $f(w,z) = 0$, que es un polinomio irreductible en w y en z, y hemos señalado lo que es una superficie de Riemann. Esta no fue la manera de hacerlo Riemann. El empezó con una superficie de Riemann y propuso mostrar que existe una ecuación $f(w,z) = 0$ que pertenece a ella y, además, que existen otras funciones univalentes y multivalentes definidas sobre la superficie de Riemann.

La definición de Riemann de una función analítica univalente $f(z) = u + iv$ es que la función es analítica en un punto y en su entorno si es continua y diferenciable y satisface lo que ahora llamamos las ecuaciones de Cauchy-Riemann

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{y} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} . \tag{46}$$

Estas ecuaciones, como sabemos, aparecieron en los trabajos de D'Alembert, Euler y Cauchy. Incidentalmente, Riemann fue el primero en requerir que la existencia de la derivada dw/dz significaba que el límite $\Delta w/\Delta z$ debe ser el mismo para *toda* aproximación de $z + \Delta z$ a z. (Esta condición distingue las funciones complejas, ya que en el caso de las funciones reales $u(x,y)$, la existencia de las primeras derivadas de u para todas las direcciones de aproximación a algún (x_0,y_0) no garantiza la analiticidad.) Entonces, buscó lo que puede ser descrito como las condiciones mínimas bajo las cuales una función de

$x + iy$ pueda ser determinada como un todo en cualquier dominio en que existiera. Es evidente a partir de las ecuaciones Cauchy-Riemann que u y v satisfacen la ecuación del potencial bidimensional

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0. \tag{47}$$

Riemann tuvo la idea de que una función compleja podía estar determinada inmediatamente en la totalidad del dominio de su existencia usando el hecho de que u satisface la ecuación del potencial.

Específicamente, Riemann supone que una función w de posición sobre la superficie de Riemann está determinada, excepto una constante aditiva, por la función real $u(x,y)$ si u está sujeta a las siguientes condiciones:

1. Satisface la ecuación del potencial en todos los puntos sobre la superficie donde sus derivadas no son infinitas.
2. Si u debe ser una función multivaluada, sus valores en cualquier punto sobre la superficie difieren en combinaciones lineales de múltiplos enteros de constantes reales. (Estas constantes reales son las partes reales de los módulos de periodicidad de w, que discutiremos más adelante.)
3. u puede tener (polos) infinitos específicos de cierta forma en puntos fijados sobre la superficie. Estos infinitos deben pertenecer a las partes reales de los términos que dan los infinitos en w. Supone, más adelante, como una condición subsidiaria el que u pueda tener valores finitos a lo largo de una curva cerrada delimitando una porción de la superficie o que haya una relación entre los valores en la frontera de u y v. Riemann es vago en cuanto a lo general que pueda ser esta relación.

Estas condiciones deben determinar u. Una vez que u está determinada, entonces, a la vista de las ecuaciones de Cauchy-Riemann,

$$v = \int \left(-\frac{\partial u}{\partial y}\, dx + \frac{\partial u}{\partial x}\, dy \right). \tag{48}$$

Así v está también determinada, y también lo está w. Es importante notar que, para Riemann, el dominio de u era cualquier parte de la superficie de Riemann, incluyendo posiblemente la superficie completa. En su tesis doctoral consideró superficies con borde y única-

mente después utilizó superficies cerradas, esto es, superficies sin fronteras, tales como un toro.

Para determinar u, la herramienta esencial de Riemann fue lo que él llamó el principio de Dirichlet, ya que lo aprendió de Dirichlet; pero lo extendió a dominios sobre las superficies de Riemann y, además, prescribió singularidades para u en el dominio y también saltos (condiciones 2 y 3 de las anteriores). El principio de Dirichlet dice que una función u que minimiza la integral de Dirichlet

$$\iint \left\{ \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right\} dx\, dy$$

satisface la ecuación del potencial. La última es, de hecho, la condición necesaria para que la primera variación de la integral de Dirichlet se anule (véase también cap. 28, sec. 4). Ya que el integrando en la integral de Dirichlet es positivo y por lo mismo tiene una cota inferior que es positiva o en el peor caso cero, Riemann concluyó que debe existir una función u que minimiza la integral y por tanto satisface la ecuación del potencial. Así, la existencia de la función u y, por tanto, mediante (48), de $f(z)$ que pertenece a la superficie de Riemann y que aún podía tener singularidades y saltos complejos (módulos de periodicidad) prescritos estaba asegurada en cuanto a Riemann se refiere.

Una vez que la existencia de las funciones sobre una superficie de Riemann, como dominio, había sido establecida, era posible demostrar que existe una ecuación fundamental que está asociada a la superficie dada; esto es, existe una $f(w,z) = 0$ que tiene a la superficie dada como su superficie. De qué manera la superficie corresponde a la relación entre w y z, Riemann no lo aclara. De hecho, esta $f(w,z) = 0$ no es única: a partir de cada función racional w_1 de w y z sobre la superficie puede obtenerse a través de $f(w,z) = 0$ otra ecuación $f_1(w_1,z) = 0$, la cual, si es irreducible, tiene la misma superficie de Riemann. Esta es una característica del método de Riemann.

Para investigar más a fondo las clases de funciones que pueden existir sobre una superficie de Riemann, es necesario conocer la noción de Riemann de la conexión de una superficie de Riemann. Una superficie de Riemann puede tener curvas frontera, o ser cerrada como una esfera o un toro. Si es la superficie riemanniana de una función algebraica, esto es si $f(w,z) = 0$ define w como una función

de z y si f es un polinomio en w y z, entonces la superficie es cerrada. Si f es irreductible, esto es, no puede ser expresado como producto de tales polinomios, entonces la superficie consiste en una pieza y se dice que es conexa.

Un plano (o una esfera) es una superficie tal que cualquier curva cerrada la divide en dos partes, de tal manera que no es posible pasar continuamente de un punto en la primera parte a un punto en la segunda sin cruzar la curva cerrada. Se dice que tal superficie es simplemente conexa. Sin embargo, si es posible dibujar alguna curva cerrada sobre la superficie y la curva no desconecta la superficie, entonces la superficie no es simplemente conexa. Por ejemplo, se pueden dibujar dos curvas cerradas diferentes sobre el toro (fig. 27.7), y ni siquiera la presencia de ambas desconecta la superficie.

Riemann deseaba asignar un número que indicara la conexión de la superficie. Consideraba los polos y los puntos de ramificación como partes de la superficie, y dado que pensaba en funciones algebraicas, sus superficies eran cerradas. Quitando una pequeña porción de una de las hojas, la superficie tenía una curva frontera C. Entonces imaginó la superficie como cortada por una curva que no se intersectaba a sí misma y que va del borde C a otro punto del borde C. Tal curva es llamada sección cruzada *(Querschnitt).* Esta sección y C son consideradas como una nueva frontera y puede ser hecho un segundo corte, empezando en algún punto de la (nueva) frontera y finalizando en otro, no cruzando la (nueva) frontera.

Un número suficiente de estas secciones cruzadas es introducido de tal forma que se corte una superficie de Riemann que pueda ser

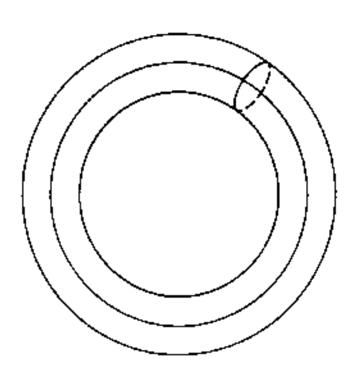


FIGURA 27.7

múltiplemente conexa en una única superficie simplemente conexa. Así, si una superficie es simplemente conexa, no se necesita lo anterior y la superficie es conexa *(Grundzahl).* Una superficie es doblemente conexa si mediante un corte apropiado es cambiada en una única superficie simplemente conexa. Entonces la conectividad es dos. Un anillo plano y una superficie esférica con dos agujeros son ejemplos de lo dicho. Una superfície es llamada triplemente conexa cuando mediante dos cortes apropiados es convertida en una única superficie simplemente conexa. Entonces la conectividad es tres. Un ejemplo es la (superficie de un) toro con un agujero. En general, se dice que una superficie es N-mente conexa o tiene conectividad (o conexión) N si mediante $N - 1$ cortes puede ser convertida en una superficie simplemente conexa. Una superficie esférica con N agujeros tiene conectividad N.

Ahora es posible relacionar la conectividad de una superficie de Riemann (con una frontera) y el número de puntos de ramificación. Cada punto de éstos, digamos, r_i, debe ser contado de acuerdo con la multiplicidad del número de ramas de la función que se cruzan en dicho punto. Si el número es w_i, $i = 1, 2, ..., r$, entonces la multiplicidad de r_i es $w_i - 1$. Supongamos que la superficie tiene q hojas. Entonces la conectividad N está dada por

$$N = \sum_i w_i - 2q + 3.$$

Se puede demostrar que la conectividad N de una superficie cerrada con una frontera singular es $2p + 1$. De aquí que

$$2p = \sum w_i - 2q + 2.$$

El entero p es llamado género de la superficie de Riemann y de la ecuación asociada $f(w,z) = 0$. Esta relación fue establecida por Riemann.

Un caso especial de considerable importancia es la superficie

$$w^2 = (z - a_1)(z - a_2) \dots (z - a_n),$$

que tiene una superficie de Riemann con dos hojas. Hay un número finito n de puntos de ramificación, y $z = \infty$ es un punto de ramifi-

cación donde n es impar. Entonces $\Sigma w_i = n$ o $n + 1$ y $2q = 4$. El género p de la superficie está dado por

$$p = \begin{cases} \dfrac{n-2}{2} & \text{cuando } n \text{ es par} \\[2ex] \dfrac{n-1}{2} & \text{cuando } n \text{ es impar.} \end{cases}$$

Dada una superficie de Riemann determinada por $f(w,z) = 0$, sabemos que w es una función unívoca de los puntos sobre la superficie. Entonces, cada función racional de w y z es también una función univaluada de posición sobre la superficie (ya que podemos reemplazar w en la función racional por su valor en términos de z). También los puntos de ramificación de esta función racional, aunque no sus polos, son los mismos que los de f. Inversamente, es posible demostrar que toda función univaluada de la posición sobre la superficie —teniendo polos de orden finito— es una función racional de w y z.

Aun en el caso de funciones univaluadas definidas sobre el plano ordinario, las integrales de tales funciones pueden ser multivaluadas. De esta manera

$$\int_0^z \frac{dz}{1 + z^2} = w + n\pi,$$

donde w es el valor de la integral a lo largo, digamos, de una trayectoria en línea recta de 0 a z, y n depende de cómo la trayectoria de 0 a z circunscriba $\pm i$. De la misma manera, cuando se consideran funciones univaluadas sobre la superficie de Riemann, digamos una función racional de w y z sobre la superficie, la integral de dicha función puede ser multivaluada; esto puede suceder. Si uno entonces introduce cortes para hacer la superficie simplemente conexa, y si la integral sigue una trayectoria de z_1 a z_2, cada vez que la trayectoria cruza un corte se suma un valor constante I al valor básico U de la integral para una trayectoria sobre una porción simplemente conexa de la superficie. Si la trayectoria debe cruzar el corte m veces en la misma dirección, entonces se suma el valor mI a U. La constante I se llama un módulo de periodicidad. Cada corte introduce sus propios módulos de periodicidad, y si la conectividad de la superficie es $N + 1$, existen N módulos de periodicidad linealmente indepen-

dientes. Sean éstos I_1, I_2, ..., I_n. Entonces el valor de la integral de la función univalente tomada sobre su trayectoria original es

$$U + m_1 I_1 + m_2 I_2 + ... + m_n I_n,$$

donde los m_1, m_2, ..., m_n son enteros. Los I_i son, generalmente, números complejos.

9. Integrales y funciones abelianas

A pesar de que cuatro importantes artículos de Riemann publicados en el *Journal fur Mathematil*[40] repiten muchas de las ideas de su disertación, están principalmente dedicados a las integrales y funciones abelianas. El cuarto ensayo es el que proporcionó a la materia su mayor desarrollo. Los cuatro son difíciles de entender: «eran libros con siete sellos». Afortunadamente, muchos matemáticos brillantes trabajaron más adelante y dieron cabal explicación del material. Riemann colocó a la par el trabajo de Abel y Jacobi, el cual surgió en gran parte de las funciones reales, y el tratamiento de Weierstrass, que se valió de funciones complejas.

A partir del hecho de que Riemann había aclarado el concepto de funciones multivaluadas, podría haber sido más claro en lo concerniente a las integrales abelianas. Sea $f(w,z) = 0$ la ecuación de una superficie de Riemann y sea $\int R(w,z,)dz$ una integral de una función racional de w y z sobre esta superficie de Riemann. Riemann clasificó las integrales abelianas como sigue. Entre las *integrales* de funciones racionales de w y z sobre una superficie de Riemann determinada por la ecuación $f(w,z) = 0$, existen algunas que, a pesar de ser funciones multivaluadas sobre la superficie sin cortar, son finitas en todas partes. Estas son llamadas integrales de primera clase. El número de tales integrales linealmente independientes es igual al género p de la superficie si la conectividad es $2p + 1$. Si se hacen $2p$ cortes, cada integral es una función univalente para una trayectoria en la región acotada por los cortes. Si la trayectoria debe cruzar un corte, entonces el módulo de periodicidad que discutimos en la sección anterior debe ser usado y el valor de la integral es: si W es su

40 Vol. 54, 1857, 115-155 = *Werke*, 88-144.

valor desde un punto fijo hasta z, entonces todos los posibles valores son

$$W = \sum_{r=1}^{2p} m_r \omega_r,$$

donde los m_r son enteros y las ω_r son módulos de periodicidad para esta integral.

Las integrales de segunda clase tienen infinitos algebraicos pero no logarítmicos. Una integral elemental de segunda clase tiene un infinito de primer orden en un punto sobre la superficie de Riemann. Si $E(z)$ es un valor de la integral en un punto sobre la superficie (el límite superior de la integral), entonces todos los valores de la integral están incluidos en

$$E(z) + \sum_{r=1}^{2p} n_r \varepsilon_r,$$

donde los n_r son enteros y las ε_r son los módulos de periodicidad para esta integral. Dos integrales elementales con un infinito en un punto en común sobre la superficie de Riemann difieren en una integral de primera clase. De aquí podemos inferir que existen $p + 1$ integrales elementales linealmente independientes de segunda clase con un infinito en el mismo punto sobre la superficie de Riemann.

Las integrales que tienen infinitos logarítmicos son llamadas integrales de tercera clase. Sucede que cada una debe tener dos infinitos logarítmicos. Si dicha integral no tiene infinitos algebraicos, esto es, no hay términos algebraicos en la expansión de la integral cerca de cualquiera de los puntos en los cuales tiene infinitos logarítmicos, entonces la integral es llamada integral *elemental* de tercera clase. Existen $p + 1$ integrales elementales linealmente independientes de tercera clase teniendo sus infinitos logarítmicos en los mismos dos puntos sobre la superficie de Riemann. Toda integral abeliana es la suma de integrales de las tres clases.

El análisis de las integrales abelianas arroja luz sobre qué clase de funciones pueden existir en una superficie de Riemann. Riemann trata dos clases de funciones; la primera consiste en funciones univaluadas sobre la superficie cuyas singularidades son polos. La segunda clase consiste en funciones que son univaluadas sobre la superficie obtenida con cortes, pero discontinuas a lo largo de cada

corte. Por supuesto, tal función difiere en una constante compleja h_v en un lado del v-ésimo corte de su valor en el otro lado. Este segundo tipo de función también puede tener polos e infinitos logarítmicos. Riemann muestra que las funciones de primera clase son algebraicas y las de segunda son integrales de funciones algebraicas.

También hay funciones que son finitas en todos los lugares sobre la superficie. Se representan tales funciones por medio de funciones de la primera clase anterior. También podemos construir funciones algebraicas sobre una superficie combinando integrales de segunda y tercera clase. De esta manera, Riemann demostró que las funciones algebraicas pueden ser representadas por una suma de funciones trascendentes. También las funciones univalentes algebraicamente infinitas en un número dado de puntos pueden representarse mediante funciones racionales. Una función que es univalente sobre la superficie entera es el integrando de una integral finita en todas partes. La función puede entonces ser representada como una función racional en w y x y puede tener la forma $\phi(w, z)/\partial f/\partial w$ donde $f(w,z) = 0$ es la ecuación de la superficie. La función ϕ que aparece aquí y en la construcción de integrales de primera clase es llamada el polinomio adjunto de $f(w,z) = 0$. Su grado es generalmente $n - 3$, cuando el grado de f es n.

La importancia de las funciones racionales sobre una superficie de Riemann se deriva del hecho, arriba mencionado, de que toda función univalente sobre la superficie y que no tenga singularidades esenciales es una función racional. Tal función tiene tantos ceros como polos y toma todo valor el mismo número de veces. Más aún, una vez que la función $f(w,z) = 0$, que define la superficie, está fijada, todas las otras funciones de posición sobre la superficie son coextensivas en su totalidad con funciones racionales de w y z e integrales de tales funciones.

Weierstrass también trabajó sobre las integrales abelianas durante la década de los 60. Pero él y los otros seguidores de Riemann en este campo construyeron funciones trascendentes a partir de funciones algebraicas, procedimiento opuesto del de Riemann. Lo hicieron así porque tenían razones para desconfiar del principio de Dirichlet. Weierstrass, en un artículo leído en 1870 [41], señaló que la existencia de una función que minimiza la integral de Dirichlet no había sido establecida. El propio Riemann tenía otra manera de pensar. Reco-

[41] *Werke*, 2, 49-54.

noció el problema de establecer la existencia de una función minimizadora para la integral de Dirichlet, antes que Weierstrass hiciera su argumento, pero declaró que el principio de Dirichlet era una herramienta conveniente que estaba disponible; la existencia de la función *u*, dijo, era correcta, de todos modos. La observación de Helmholtz en este punto es también interesante: «... para los físicos el [uso del] principio de Dirichlet se mantiene como prueba» [42].

Otra de las nuevas investigaciones en teoría de funciones complejas que inició Riemann es la inversión de las integrales abelianas, esto es, determinar la función z de *u* cuando

$$u = \int_0^z R(z, w)\, dz$$

y, por supuesto, *w* y *z* están relacionadas por una ecuación algebraica. La función z de *u* no es únicamente multivaluada; peor, no es claramente definible. Como en el caso de las integrales hiperelípticas, Riemann tomó sumas de *p* integrales abelianas y definió nuevas *funciones abelianas* de *p* variables que son univalentes y 2*p*-múltiplemente periódicas. Por una función 2*p*-multiplemente periódica de *p* variables se entiende que existen 2*p* conjuntos de cantidades w_{1k}, w_{2k}, ..., w_{pk} $k = 1, 2, \ldots, 2p$, conteniendo cada conjunto un período de cada una de las *p* variables. Riemann demostró que una función univalente no puede tener más de dos 2*p* conjuntos de períodos simultáneos. Las funciones abelianas, expresadas en términos de funciones theta en *p* variables, son generalizaciones de las funciones elípticas.

Uno de los resultados más notables para las funciones sobre superficies de Riemann de género *p* es conocido ahora como el teorema de Riemann-Roch. El trabajo sobre este resultado fue iniciado por Riemann y completado por Gustav Roch (1839-1866) [43]. Esencialmente, el teorema determina el número de funciones meromorfas linealmente independientes sobre la superficie que tienen a lo más un conjunto de específico finito de polos. Más detalladamente, supongamos que *w* es una función univaluada sobre la superficie y tiene polos de primer orden en los puntos $c_1, \ldots, c_m$ pero no en otro

---

[42] Para la historia posterior del problema de Dirichlet y el principio de Dirchlet, véase Cap. 28, secs. 4 y 8.

[43] *Jour. für Math.*, 64, 1864, 372-376.

lado. Las posiciones c_i no son necesariamente independientes. Si q funciones linealmente independientes (funciones adjuntas) se anulan en ellos, entonces w contiene $m - p + q + 1$ constantes arbitrarias, y es una combinación lineal de múltiplos arbitrarios de $m - p + q$ funciones, cada una con $p - q + 1$ polos del primer orden, $p - q$ de los cuales son comunes a todas las funciones en la combinación.

10. Aplicaciones conformes

Para completar la teoría de su tesis doctoral, Riemann acaba con algunas aplicaciones de la teoría de funciones a la representación conforme. El problema general de la representación conforme de un plano en un plano (que proviene del trazado de mapas) fue resuelto por Gauss en 1825. Su resultado se reduce al hecho que tal aplicación es establecida por cualquier $f(z)$ analítica —aunque Gauss no utilizó la teoría de funciones complejas—. Riemann sabía que una función analítica establecía una representación conforme del plano z en el plano w, pero le interesaba extender esto a las superficies de Riemann. Así se abrió un nuevo capítulo en la representación conforme.

Al final de su tesis Riemann proporciona el siguiente teorema: dadas dos superficies planas simplemente conexas (incluye dominios simplemente conexos sobre superficies de Riemann) pueden ser aplicadas uno a uno y conformemente una sobre otra, y un punto interior y un punto frontera sobre una superficie pueden ser asignados a puntos interiores y puntos frontera sobre el otro escogidos arbitrariamente. Así, la aplicación está determinada en su totalidad. Este teorema contiene como caso especial el resultado básico de que, dado cualquier dominio simplemente conexo D con una frontera que contiene más que un punto y dado un punto A de este dominio y una dirección T en este punto, existe una función $w = f(z)$ que es analítica en D y aplica D conforme y biunívocamente dentro de un círculo de radio 1 centrado en el origen del plano w. Bajo esta aplicación A va al origen y T es enviado en la dirección del eje real positivo. Este último aserto es descrito usualmente como el teorema de la aplicación de Riemann.

Riemann demostró este teorema usando el principio de Dirichlet, pero en aquellos días, al hallarse que era erróneo, los matemáticos buscaron una prueba más sólida. Carl Gottfried Neumann y Her-

mann Amandus Schwarz demostraron (1870) que era posible aplicar una región plana simplemente conexa sobre un círculo. Sin embargo, no pudieron manejar dominios simplemente conexos con varias hojas.

En algunas ocasiones, el interés dedicado a aplicar una región simplemente conexa conformemente sobre un círculo es explicado por el hecho de que, para aplicar una región simplemente conexa sobre otra conformemente, es suficiente aplicar cada una sobre un círculo y entonces el producto de las dos aplicaciones conformes hará el resto.

Mientras que la prueba del teorema de Riemann permanecía abierta, se obtuvieron varios resultados especiales sobre aplicaciones conformes. De éstos, uno de los más útiles para la solución de ecuaciones en derivadas parciales fue dado por Schwarz [44] y Elwin Bruno Christoffel [45]. Su teorema muestra cómo aplicar un polígono y su interior (Fig. 27.8) en el plano z, conformemente, dentro de la mitad superior del plano w. La aplicación está dada por

$$z = c\int_a^z (w-a)^{(\alpha/\pi)-1}(w-b)^{(\beta/\pi)-1}\dots dw + c',$$

donde c y c' son determinables a partir de la posición del polígono y donde a, b, c,..., corresponden a A, B, C,... Esta aplicación ha probado ser muy útil para resolver la ecuación del potencial (de Laplace).

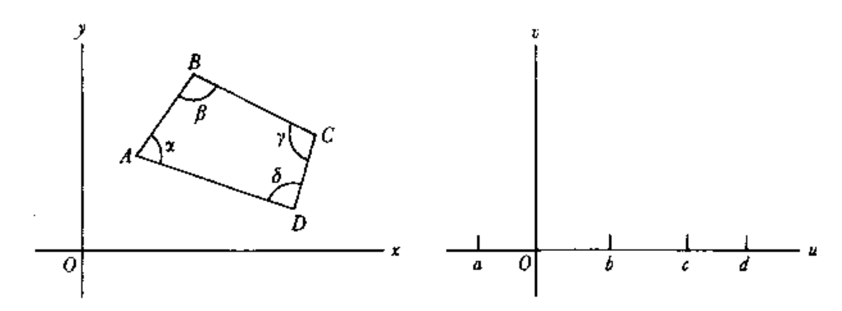


FIGURA 27.8

[44] *Jour. für Math.*, 70, 1869, 105-120 = *Ges. Abh.*, 2, 65-83.

[45] *Annali di Mat.*, (2), 1, 1867, 95-103, y (2), 4, 1871, 1-9 = *Ges. Abh.*, 2, 56 sgs.

11. La representación de funciones y los valores excepcionales

El desarrollo de la teoría de funciones complejas continuó en la segunda mitad del siglo XIX y tendremos ocasión, en capítulos posteriores, de considerar algunos resultados. Sin embargo, unas pocas de entre tantas creaciones que se apoyan primordialmente sobre las propias funciones complejas serán tratadas aquí.

Entre las funciones complejas univalentes, las funciones enteras, esto es, aquellas que no tienen singularidades en la parte finita del plano, que incluyen polinomios, e^z, sen z y cos z demostró ser de considerable interés, ya que son, *groso modo*, las análogas de las funciones reales elementales. Para tales funciones, el teorema de Liouville establece que toda función entera acotada es una constante [46]. Weierstrass extendió a funciones enteras el teorema sobre descomposición de polinomios reales en factores lineales. El teorema de Weierstrass [47], el cual probablemente proviene de los 40, es llamado teorema de factorización, y dice que si $G(z)$ es una función entera que no se anula idénticamente sino que tiene un número infinito de raíces (esto es, no es un polinomio), entonces $G(z)$ puede ser escrita como un producto infinito

$$G(z) = \Gamma(z) z^m \prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right) e^{g_n(z)},$$

donde

$$g_n(z) = \frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \dots + \frac{1}{m_n}\left(\frac{z}{a_n}\right)^{m_n}.$$

$\Gamma(z)$ es una función entera sin ceros; las a_n son los ceros de $G(z)$ y z^m representa el cero en $z = 0$ de multiplicidad m si $G(z)$ tiene ese cero. Los factores de los productos son llamados factores primos de $G(z)$.

A las funciones enteras siguen en complejidad las funciones meromorfas, las cuales pueden tener únicamente polos en la región finita del plano complejo. En su ensayo de 1876 [48], Weierstrass de-

[46] El teorema se le debe a Cauchy (*Comp. Rend.*, 19, 1844, 1377-1381 = *Œuvres*, (1), 8, 378-385). C. W. Borchardt lo oyó en las clases de Liouville de 1847 y a él atribuyó el teorema.

[47] *Abh. König. Akad. der Wiss., Berlin*, 1876, 11-60 = *Werke*, 2, 77-124.

[48] *Abh. König. Akad. der Wiss., Berlin*, 1876, 11-60 = *Werke*, 2, 77-124.

mostró que una función meromorfa puede ser expresada como un cociente de dos funciones enteras. El teorema fue extendido por Gösta Mittag-Leffler (1846-1927) en un ensayo de 1877 [49]. Una función meromorfa en una región arbitraria puede ser expresada como cociente de dos funciones, ambas analíticas en la región. En ambos teoremas de Weierstrass y Mittag-Leffler, el numerador y denominador no se anulan en el mismo lugar en la región.

Otro tema que llamó la atención de numerosos matemáticos es el del rango de valores que pueden tomar diversos tipos de funciones complejas. (Charles) Emile Picard (1856-1941), profesor de análisis superior en la Sorbona y secretario perpetuo de la Academia de Ciencias de París, obtuvo una serie de resultados. En 1879 [50], Picard demostró que una función entera puede omitir a lo más un valor finito sin reducirse a una constante, y si existieran al menos dos valores cada uno de los cuales es tomado un número finito de veces, entonces la función es polinómica. En cualquier otro caso, la función toma cada valor, aparte del excepcional, un número infinito de veces. Si la función es meromorfa, siendo el infinito un valor admisible, pueden ser omitidos a lo más dos valores sin que la función se reduzca a una constante.

En el mismo artículo, extendió un resultado de Julian W. Sochozki (1842-1927) y Weierstrass, y demostró que en cualquier entorno de un punto singular esencial aislado, una función toma todos los valores, a excepción de a lo más un valor (finito). El resultado es muy profundo y tiene multitud de consecuencias. Por supuesto, un buen número de otros resultados y pruebas alternativas fueron creados y llevaron el problema hasta bien entrado el siglo XX.

En el campo de las funciones complejas, el siglo XIX finalizó con el regreso a los fundamentos. Las demostraciones del teorema de la integral de Cauchy en el siglo XIX se basaron en el hecho de que df/dz es continua. Edouard Goursat (1858-1936) demostró [51] el teorema de Cauchy, $\int f(z)dz = 0$ alrededor de una curva cerrada C, sin suponer la continuidad de la derivada $f'(z)$ en la región cerrada limitada por la curva C. La existencia de $f'(z)$ era suficiente. Goursat señaló que la continuidad de $f(z)$ y la existencia de la derivada eran suficientes para caracterizar la analiticidad.

---

[49] *Öfversight of Kongliga Vetenskops-Akademiens Förhandlingar*, 34, 1877, 1, 17-43; véase también *Acta Math.*, 4, 1884, 1-79.

[50] *Ann. de l'Ecole Norm. Sup.*, (2), 9, 1880, 145-166.

[51] *Amer. Math. Soc. Trans.*, 1, 1900, 14-16.

Como ha mostrado nuestro bosquejo del surgimiento de la teoría de las funciones complejas, queda demostrado que Cauchy, Riemann y Weierstrass son los tres principales fundadores de la teoría de funciones. Por largo tiempo, sus respectivas ideas, y métodos, fueron seguidos independientemente por sus seguidores. Más adelante se fusionaron las ideas de Cauchy y Riemann y las ideas de Weierstrass fueron gradualmente reducidas a partir del punto de vista Cauchy-Riemann, de tal forma que la idea de empezar a partir de las series de potencias ya no es básica. El rigor de la visión de Cauchy-Riemann se mejoró, de tal forma que desde este punto de vista la perspectiva de Weierstrass no se toma como esencial. La unificación completa se llevó a cabo al principio del siglo XX.

Bibliografía

Abel, N. H.: *Œuvres complètes,* 2 vols., 1881, Johnson Reprint Corp., 1964.

— : *Mémorial publié à l'occasion du centénaire de sa naissance,* Jacob Dibwad, Cristianía, 1902.

Brill, A., y Noether, M.: «Die Entwicklung der Theorie der algebraischen Funktionen in älterer und neuerer Zeit», *Jahres. der Deut. Math. -Verein.,* 3, 1892/1893, 109-556, en particular 155-186.

Brun, Viggo: «Niels Henrik Abel. Neue biographische Funde». *Jour. fur Math.,* 193. 1954, 239-249.

Cauchy, A. L.: *Œuvres complètes,* 26 vols., Gauthier-Villars, 1882-1938, ensayos relevantes.

Crowe, Michael J.: *A history of vector analysis,* University of Notre Dame Press, 1967, Cap. I.

Enneper, A.: *Elliptische Funktionen: Theorie und Geschichte,* 2.ª ed., L. Nebert, 1890.

Hadamard, Jacques: *Notice sur les travaux scientifiques de M. Jacques Hadamard,* Gauthier-Villars, 1901.

Jacobi, C. G. J.: *Gesammelte Werke,* 7 vols., y suplemento, G. Reimer, 1881-1891; Chelsea reprint, 1968.

Jourdain, Philip E. B.: «The Theory of Functions with Cauchy and Gauss», *Bibliotecha Mathematica,* (3), 6, 1905, 190-207.

Klein, Felix: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert,* 2 vols., Chelsea (reimpresión), 1950.

Levy, Paul, et al.: «La vie et l'œuvre de J. Hadamard», *L'Enseignement Mathématique,* (2), 13, 1967, 1-72.

Markuschewitsch, A. I.: *Skizzen zur Geschichte der analytischen Funktionen,* V. E. B. Deutscher Verlag der Wissenschaften, 1955.

Mittag-Leffler, G.: «An introduction to the Theory of Elliptic Functions», *Annals of Math.*, 24, 1922-1923, 271-351.
— :«Die ersten 40 Jahre des Lebens von Weierstrass», *Acta Math.*, 39, 1923, 1-57.
Ore, O.: *Niels Henrik Abel, Mathematician Extraordinary*, University of Minnesota Press, 1957.
Osgood, W. F.: «Allgemeine Theorie der analytischen Funktionen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1901-1921, II B1, 1-114.
Reichardt, Hans, ed.: *Gauss: Leben und Werke*, Haude und Spenersche Verlagsbuchhandlung, 1960; G. Teubner, 1957, 151-182.
Riemann, Berhard: *Gesammelte mathematische Werke*, 2.ª ed., Dover (reimpresión), 1953.
Schlesinger, L.: «Über Gauss' Arbeiten zu Funktionenlehre», *Nachrichten König. Ges. der Wiss. zu Gött.*, 1912, Beiheft, 1-143, también en los *Werke* de Gauss, 10, 77 sgs.
Smith, David Eugene: *A Source Book in Mathematics.*, 2 vols., Dover (reimpresión), 1959, pp. 55-66, 404-410.
Staeckel, Paul: «Integration durch imnaginäres Gebiet», *Bibliotecha Mathematica*, (3), 1, 1900, 109-128.
Valson, C. A.: *La vie et les travaux du baron Cauchy*, 2 vols., Gauthier-Villars, 1868.
Weierstrass, Karl: *Mathematische Werke*, 7 vols., Mayer und Müller, 1895-1924.

Capítulo 28
LAS ECUACIONES EN DERIVADAS PARCIALES EN EL SIGLO XIX

El estudio profundo de la naturaleza es la fuente más fértil de los descubrimientos matemáticos.

JOSEPH FOURIER

1. Introducción

El campo de las ecuaciones en derivadas parciales, que tuvo sus inicios en el siglo XVIII, rejuveneció en el XIX. De la misma manera que se extendieron las ciencias naturales, tanto en la variedad como en la profundidad de los fenómenos investigados, el número de nuevos tipos de ecuaciones diferenciales se incrementó; y aun los tipos ya conocidos, la ecuación de ondas y la ecuación del potencial, fueron aplicados a nuevas áreas de la física. Las ecuaciones en derivadas parciales se convirtieron, y permanecieron, en el corazón de las matemáticas. Su importancia para las ciencias físicas es únicamente una de las razones para asignarles este lugar central. Desde el punto de vista de las propias matemáticas, la solución de ecuaciones diferenciales parciales creó la necesidad de desarrollos matemáticos en la teoría de funciones, el cálculo de variaciones, desarrollos en series, ecuaciones diferenciales ordinarias, álgebra y geometría diferencial. La materia ha logrado tal extension que en este capítulo únicamente podemos describir algunos de los resultados más importantes.

Hoy en día, estamos acostumbrados a clasificar las ecuaciones diferenciales en tipos. Al inicio del siglo XIX, se sabía tan poco de

la materia que no se podía haber tenido la idea de distinguir los distintos tipos. Los problemas físicos dictaban qué ecuaciones debían ser atacadas y los matemáticos pasaban libremente de un tipo a otro sin reconocer diferencias entre ellos, que hoy en día consideramos fundamentales. El mundo físico era, y es aún, indiferente a las clasificaciones matemáticas.

2. La ecuación de calor y las series de Fourier

El primer gran paso del siglo XIX, y desde luego uno de enorme importancia, fue dado por Joseph Fourier (1768-1830). Fourier había sido un muy buen estudiante de matemáticas, pero se había propuesto convertirse en un oficial del ejército. Negándosele una comisión, ya que era hijo de un sastre, se refugió en el sacerdocio. Cuando se le ofreció una plaza de profesor en la escuela militar a la que él había asistido aceptó y entonces las matemáticas se covirtieron en el interés de su vida.

Como otros científicos en sus días, Fourier se ocupó del flujo del calor. El flujo tenía interés como problema práctico en el manejo de metales en la industria, y como problema científico en los intentos por determinar la temperatura en el interior de la tierra, la variación de dicha temperatura con respecto al tiempo, y otras cuestiones. Sometió un ensayo básico sobre la conducción del calor a la Academia de Ciencias de París en 1807 [1]. El artículo fue juzgado por Lagrange, Laplace y Legendre y fue rechazado. Pero la Academia deseaba motivar a Fourier para desarrollar sus ideas y propuso el problema de la propagación del calor como materia del gran premio que sería asignado en 1812. Fourier sometió una versión revisada en 1811, que fue juzgada por los anteriormente mencionados, y otros. Ganó el premio, pero fue criticado por su falta de rigor y por lo mismo no fue publicado en las *Mémoires* de la Academia en aquellos días. Fourier se resintió del trato de que fue objeto. Continuó trabajando sobre el calor y, en 1822, publicó con uno de los clásicos de las matemáticas: *Théorie Analytique de la chaleur (Teoría analítica del calor)* [2]. Incorporaba la primera parte de su artículo de 1811, prácticamente sin un solo cambio. Este libro es la fuente prin-

[1] El manuscrito se encuentra en la *Ecole des Ponts et Chaussées*.

[2] *Œuvres*, 1.

cipal para las ideas de Fourier. Dos años más tarde se convirtió en el secretario de la Academia y vio la oportunidad de hacer que se publicara en las *Mémoires*[3] su artículo de 1811 conservando su forma original.

En el interior de un cuerpo sujeto a pérdida o aumento de calor, por lo general la temperatura no está distribuida uniformemente y cambia en cualquier lugar con el tiempo, lo que hace que la temperatura T sea una función del espacio y el tiempo. La forma precisa de la función depende del contorno del cuerpo, la densidad, el calor específico del material, la distribución inicial de T, esto es, la distribución en el tiempo $t = 0$, y las condiciones mantenidas sobre la superficie del cuerpo. El primer problema importante que Fourier consideró en su libro fue la determinación de la temperatura T en un cuerpo homogéneo e isótropo como una función de x, y, z y t. Demostró, sobre la base de principios físicos, que T debe satisfacer la ecuación diferencial parcial, llamada ecuación de calor en tres dimensiones,

$$\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) = k^2 \frac{\partial T}{\partial t}, \tag{1}$$

donde k^2 es una constante cuyo valor depende del material del cuerpo.

Más adelante, Fourier resolvió problemas específicos de conducción de calor. Debemos considerar un caso que es típico de su método, el problema de resolver la ecuación (1) para la barra cilíndrica cuyos extremos se mantienen a 0° de temperatura y cuya superficie lateral está aislada de tal manera que el calor no fluye a través de ella. Ya que esta barra supone únicamente un espacio de una dimensión (1) se convierte en

$$\frac{\partial^2 T}{\partial x^2} = k^2 \frac{\partial T}{\partial t} \tag{2}$$

sujeta a las condiciones de contorno

$$T(0, t) = 0, \quad T(l, t) = 0, \quad \text{para } t > 0, \tag{3}$$

[3] *Mém. de l'Acad. des Sci., Paris* (2), 4. 1819-1820, 185-555. pub, 1824, y 5, 1821-1822, 153-246, pub. 1826; únicamente la segunda parte se encuentra reproducida en las *Œuvres* de Fourier, 2, 3-94.

y la condición inicial

$$T(x, 0) = f(x) \quad \text{para } 0 < x < l. \tag{4}$$

Para resolver este problema, Fourier usó el método de separación de variables. Consideró

$$T(x, t) = \phi(x)\psi(t). \tag{5}$$

Substituyendo en la ecuación diferencial, obtuvo

$$\frac{\phi''(x)}{k^2\phi(x)} = \frac{\psi'(t)}{\psi(t)}.$$

Entonces argumentó (*cf.* [30] del cap. 22) que cada una de estas razones debía ser una constante, $-\lambda$, de tal forma que

$$\phi''(x) + \lambda k^2\phi(x) = 0 \tag{6}$$

y

$$\psi'(t) + \lambda\psi(t) = 0. \tag{7}$$

Sin embargo, las condiciones (3), en virtud de (5), implicaban que

$$\phi(0) = 0 \quad \text{y} \quad \phi(l) = 0. \tag{8}$$

La solución general de (6) es

$$\phi(x) = b \operatorname{sen} (\sqrt{\lambda}kx + c).$$

La condición de que ϕ $(0) = 0$ implica que $c = 0$. La condición $\phi(1) = 0$ impone una limitación sobre λ, a saber que $\sqrt{\lambda}$ debe ser un múltiplo entero de π/kl. De aquí que haya un número infinito de valores admisibles λ_ν de λ, o bien

$$\lambda_\nu = \left(\frac{\nu\pi}{kl}\right)^2, \quad \nu \text{ entero.} \tag{9}$$

Estas λ_ν son lo que nosotros llamamos ahora los valores propios o valores característicos.

Ya que la solución general de (7) es una función exponencial y ahora λ está limitada a λ_ν, entonces, en virtud de (5), Fourier obtuvo que

$$T_\nu(x, t) = b_\nu e^{-(\nu^2\pi^2/k^2l^2)t} \operatorname{sen} \frac{\nu\pi x}{l},$$

donde la b_ν denota la constante en lugar de b y $\nu = 1, 2, 3,...$ Sin embargo, la ecuación (2) es lineal, de tal forma que una suma de soluciones es solución. De aquí se puede aseverar que

$$T(x, t) = \sum_{\nu=1}^{\infty} b_\nu e^{-(\nu^2\pi^2/k^2l^2)t} \operatorname{sen} \frac{\nu\pi x}{l}. \tag{10}$$

Para satisfacer la condición inicial (4), uno debe tener para $t = 0$

$$f(x) = \sum_{\nu=1}^{\infty} b_\nu \operatorname{sen} \frac{\nu\pi x}{l}. \tag{11}$$

Fourier se enfrentó entonces a la cuestión ¿puede $f(x)$ ser representada como una serie trigonométrica? En particular, ¿pueden ser determinadas la b_ν?

Fourier procedió a contestar estas cuestiones. Aunque para ese tiempo era ya algo consciente del problema del rigor, procedió formalmente en el espíritu del siglo XVIII. Para seguir el trabajo de Fourier haremos, para simplificar, $l = \pi$. Así consideramos

$$f(x) = \sum_{\nu=1}^{\infty} b_\nu \operatorname{sen} \nu x, \qquad \text{para } 0 < x < \pi. \tag{12}$$

Fourier toma cada función seno y la escribe mediante el teorema de Maclaurin en una serie de potencias; esto es, usa

$$\operatorname{sen} \nu x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\nu^{2n-1}}{(2n-1)!} x^{2n-1} \tag{13}$$

para reemplazar sen νx en (12). Entonces, intercambiando el orden de las sumatorias, operación incuestionada en aquel tiempo, obtiene

$$f(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!} \left(\sum_{\nu=1}^{\infty} \nu^{2n-1} b_\nu \right) x^{2n-1}. \tag{14}$$

Así $f(x)$ está expresada como una serie de potencias en x, lo que implica una grave restricción sobre la admisibilidad de $f(x)$ que no fue presupuesta para la $f(x)$ que Fourier trata. Esta serie de potencias debe ser la serie de Maclaurin para $f(x)$, de tal forma que

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0)x^k. \tag{15}$$

Igualando coeficientes de potencias similares de x en (14) y (15), Fourier encuentra que $f^{(k)}(0) = 0$ para k par y también que

$$\sum_{\nu=1}^{\infty} \nu^{2n-1} b_\nu = (-1)^{n-1} f^{(2n-1)}(0), \quad n = 1, 2, 3, \ldots .$$

Ahora, las derivadas de $f(x)$ son conocidas, ya que $f(x)$ es una condición inicial dada. De aquí que las b sean un conjunto infinito de incógnitas en un sistema infinito de ecuaciones lineales algebraicas.

En un problema previo, donde se había enfrentado a este mismo tipo de sistema, Fourier tomó los primeros k términos y la constante de la derecha de las primeras k ecuaciones, resolvió éstas, y, obteniendo una expresión general para las $b_{\nu,k}$, las cuales denotan los valores aproximados de las b_ν obtenidos a partir de las primeras k ecuaciones, concluyó audazmente que $b_\nu = \lim_{k\to\infty} b_{\nu,k}$. Sin embargo, esta vez tuvo muchas dificultades para determinar las b_ν. Tomó varias $f(x)$ diferentes y mostró cómo determinar las b_ν por procedimientos muy complicados que involucran expresiones divergentes. Usando estos casos especiales como guía obtuvo una expresión para las b_ν involucrando productos infinitos y sumas infinitas. Fourier se dio cuenta de que esta expresión no tenía en realidad utilidad, y dando más pasos audaces e ingeniosos, aunque muchas veces cuestionables, llegó finalmente a la fórmula

$$b_\nu = \frac{2}{\pi} \int_0^\pi f(s) \operatorname{sen} \nu s \, ds. \tag{16}$$

La conclusión no era, hasta cierto punto, nueva. Nosotros hemos descrito (cap. 20, sec. 5) cómo Clairut y Euler habían desarrollado algunas funciones en series de Fourier obteniendo las fórmulas

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \; b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx, \; n \geq 1. \tag{17}$$

Además, los resultados de Fourier hasta ahora derivados estaban limitados porque supuso que su $f(x)$ tenía un desarrollo de Maclaurin, y ello significa un número infinito de derivadas. Finalmente, el método de Fourier por cierto que no era riguroso y sí más complicado que el de Euler. Mientras que Fourier tuvo que usar un sistema infinito de ecuaciones lineales algebraicas, Euler procedió con mayor simplicidad empleando las propiedades de las funciones trigonométricas.

Entonces Fourier realizó algunas observaciones sorprendentes. Notó que cada b_ν podía ser interpretada como el área bajo la curva $y = (2/\pi)f(x)$ sen νx para x entre 0 y π. Tal área tiene sentido aún para funciones muy arbitrarias. Las funciones no requieren ser continuas o pueden ser conocidas sólo gráficamente. De aquí, Fourier concluyó que *toda* función $f(x)$ podía ser representada como

$$f(x) = \sum_{\nu=1}^{\infty} b_\nu \operatorname{sen} \nu x, \qquad \text{para } 0 < x < \pi. \tag{18}$$

Esta posibilidad había sido, por supuesto, rechazada por los maestros del siglo XVIII, excepto por Daniel Bernoulli.

Cuánto sabía Fourier del trabajo de sus predecesores, no está claro. En un artículo de 1825, afirma que Lacroix le había informado del trabajo de Euler, pero no dice cuándo sucedió esto. De cualquier manera, Fourier no fue disuadido por el trabajo de sus predecesores. Tomó una gran variedad de funciones $f(x)$, calculó las primeras b_ν para cada función y dibujó la suma de unos pocos primeros términos de la serie de senos (18) para cada una. A partir de esta evidencia gráfica dedujo que la serie siempre representa $f(x)$ sobre $0 < x < \pi$, sin importar si la representación se mantiene fuera del intervalo. Señala en su libro (p. 198) que dos funciones pueden coincidir en un cierto intervalo pero no necesariamente fuera del intervalo. La incapacidad de ver esto explica por qué matemáticos de épocas anteriores no podían aceptar que una función arbitraria podía ser desarrollada en una serie trigonométrica. Lo que la serie proporciona es la función en el dominio de 0 a π, en el caso presente, y repeticiones periódicas de ella fuera.

Una vez que Fourier obtuvo el sencillo resultado anterior para las b_ν, como Euler, se dio cuenta de que cada b_ν podía ser obtenida multiplicando la serie (18) por sen νx e integrando de 0 a π. También

señala que este procedimiento es aplicable a la representación

$$f(x) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty} a_\nu \cos \nu x. \tag{19}$$

Después considera la representación de cualquier $f(x)$ en el intervalo $(-\pi, \pi)$. La serie (18) representa una función impar ($f(x) = -f(-x)$) y la serie (19) una función par ($f(x) = f(-x)$). Pero cualquier función puede ser representada como la suma de una función impar $f_0(x)$ y una función par $f_e(x)$ donde

$$f_0(x) = \frac{1}{2}[f(x) - f(-x)], \quad f_e = \frac{1}{2}[f(x) + f(-x)].$$

Entonces, cualquier $f(x)$ puede ser representada en el intervalo $(-\pi, \pi)$ por

$$f(x) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu x + b_\nu \operatorname{sen} \nu x) \tag{20}$$

y los coeficientes pueden ser determinados multiplicando por $\cos \nu x$ o $\operatorname{sen} \nu x$ e integrando de $-\pi$ a π, lo que lleva a (17).

Fourier nunca proporcionó una prueba completa con respecto a que una función «arbitraria» se podía representar por una serie como (20). En su libro proporciona algunos elementos sueltos, y en la discusión final de este punto (párrafos 415, 416 y 423) ofrece un esbozo de la prueba. Pero, aun ahí, Fourier no establece las condiciones que una función debe satisfacer para ser desarrollable en serie trigonométrica. Sin embargo, la convicción de Fourier de que esto era posible está expresada a lo largo de todo el libro. También dice [4] que esta serie era convergente sin importar cómo pudiera ser $f(x)$, tanto si era posible o no asignar una expresión analítica a $f(x)$ y si la función sigue o no una ley regular. La convicción de Fourier de que cualquier función puede ser desarrollada en serie de Fourier se apoyaba sobre la evidencia geométrica descrita anteriormente. Acerca de esto dice en su libro (p. 206): «Nada nos ha parecido a nosotros más conveniente que las construcciones geométricas para de-

[4] Page 196 = *Œuvres*, 1, 210.

mostrar la verdad de los nuevos resultados y para proporcionar claramente las formas que el análisis emplea para sus expresiones.»

El trabajo de Fourier incorporó varios avances importantes. Además de llevar más lejos la teoría de las ecuaciones diferenciales parciales, forzó una revisión de la noción de función. Supóngase que la función $y = x$ está representada por una serie de Fourier (20) en el intervalo $(-\pi, \pi)$. La serie repite su comportamiento en cada intervalo de longitud 2π. De aquí que la función dada en la serie se vea como está trazada en la Figura 28.1. Tales funciones no pueden ser representadas por una expresión simple (finita) analítica, mientras que los predecesores de Fourier habían insistido en que una función debe ser representable por una única expresión. Ya que la función en su totalidad $y = x$ para toda x no está representada por la serie, no podían ver cómo una función arbitraria, que no es periódica, podía ser representada por la serie, a pesar de que Euler y Lagrange lo habían hecho así para funciones particulares no periódicas. Fourier es explícito al decir que esta serie puede representar funciones que tienen también diferentes expresiones analíticas en diferentes partes del intervalo $(0, \pi)$ o $(-\pi, \pi)$, tanto si las expresiones se unían o no continuamente. Señala, finalmente, que su trabajo clarifica los argumentos sobre las soluciones del problema de la cuerda vibrante en favor de Daniel Bernoulli. El trabajo de Fourier marcó la separación de las funciones analíticas o funciones desarrollables en series de Taylor. También es significativo que una serie de Fourier representa una función sobre un intervalo en su totalidad, mientras que una serie de Taylor representa una función únicamente en un entorno de un punto en el cual la función es analítica, aunque en casos especiales el radio de convergencia puede ser infinito.

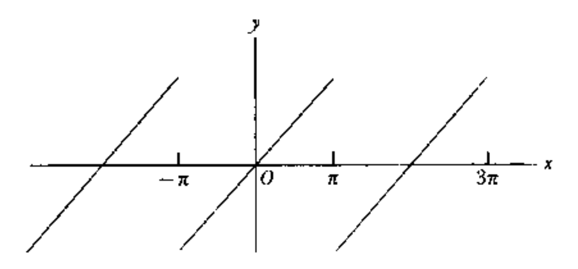


FIGURA 28.1

Ya hemos mencionado que el ensayo de 1807 de Fourier, en el que había sostenido que una función arbitraria puede ser desarrollada en una serie trigonométrica, no fue bien recibido por la Academia de Ciencias de París. En particular, Lagrange negó firmemente la posibilidad de tales desarrollos. A pesar de que únicamente criticó la falta de rigor en el artículo, estaba ciertamente muy incómodo por la generalidad de las funciones que Fourier manejaba, ya que Lagrange aún creía que una función estaba determinada por sus valores en un intervalo arbitrariamente pequeño, lo que es cierto para las funciones analíticas. De hecho, Lagrange regresó al problema de la cuerda vibrante y, sin mayor perspicacia que la que ya había mostrado en trabajos anteriores, insistió en defender la afirmación de Euler de que una función arbitraria no podía ser representada por una serie trigonométrica. Poisson aseguró posteriormente que Lagrange había mostrado que una función arbitraria puede ser representada por una serie de Fourier, pero Poisson —quien envidiaba a Fourier— dijo esto para robarle el mérito a Fourier y dárselo a Lagrange.

El trabajo de Fourier también hizo explícito otro hecho que estaba implícito en el trabajo de Euler y Laplace en el siglo XVIII. Estos habían desarrollado funciones en series de funciones de Bessel y polinomios de Legendre para poder resolver problemas específicos. El hecho general que una función podía ser desarrollada en una serie de funciones como las funciones trigonométricas, funciones de Bessel y polinomios de Legendre fue puesto a la luz por el trabajo de Fourier. Demostró, más adelante, cómo la condición inicial impuesta sobre la solución de una ecuación diferencial parcial podía ser satisfecha y así se avanzó en las técnicas de solución de dichas ecuaciones. El ensayo de Fourier de 1811, aunque no fue publicado hasta 1824-1826, mientras tanto fue hecho accesible a otros; sus ideas, en un principio aceptadas dificultosamente, finalmente ganaron su aceptación.

El método de Fourier fue seguido inmediatamente por Simeon Denis Poisson (1781-1840), uno de los más grandes analistas del siglo XIX y un físico matemático de primera clase. A pesar de que su padre había querido que estudiara medicina, se convirtió en estudiante y más tarde profesor de la fuente francesa de las matemáticas del siglo XIX: l'Ecole Polytechnique. Trabajó en la teoría del calor, fue uno de los fundadores de la teoría matemática de la elasticidad y uno de los primeros en sugerir que la teoría del potencial

gravitacional podía ser extendida a la electricidad estática y al magnetismo.

Poisson estaba tan impresionado con la afirmación de Fourier de que funciones arbitraria podían ser desarrolladas en una serie de funciones que pensó que todas las ecuaciones diferenciales parciales podían ser resueltas mediante series; cada término de la serie sería asimismo un producto de funciones (*cf.* [10]), uno para cada variable independiente. Estas expansiones, pensó, comprenden las soluciones más generales. También creyó que si una expansión divergía, eso significaba que se debía buscar una expansión en términos de otras funciones. Por supuesto, Poisson era demasiado optimista.

A partir de 1815, él mismo resolvió buen número de problemas de conducción del calor y utilizó desarrollos en funciones trigonométricas, polinomios de Legendre y armónicos superficiales de Laplace. Encontraremos parte de este trabajo más adelante. Gran parte del trabajo de Poisson sobre conducción de calor fue presentado en su *Théorie mathématique de la chaleur* (*Teoría matemática del calor*, 1835).

3. Soluciones explícitas: la integral de Fourier

A pesar del éxito e impacto de las series de Fourier como soluciones de ecuaciones diferenciales parciales, uno de los mayores esfuerzos durante el siglo XIX fue hallar soluciones en forma explícita, esto es, en términos de funciones elementales e integrales de tales funciones. Tales soluciones, al menos las de tipo conocido en los siglos XVIII y XIX, eran más manejables, más perspicaces y más fácilmente usables en los cálculos.

El método más significativo para resolver ecuaciones diferenciales parciales en forma explícita, que surgió del trabajo iniciado por Laplace, fue la integral de Fourier. La idea se debe a Fourier, Cauchy y Poisson. Es imposible asignar prioridad a este descubrimiento tan importante, ya que todos ellos presentaron ensayos orales a la Academia de Ciencias que no fueron publicados sino hasta algún tiempo después. Pero cada uno escuchó los ensayos de los otros, y resulta imposible aseverar, a partir de las publicaciones, lo que cada uno de ellos tomó de las versiones orales.

En la última sección de su ensayo ganador de 1811, Fourier trató la propagación del calor en dominios que se extendían al infinito en

una dirección. Para obtener una respuesta a tales problemas, empieza con la forma general de la solución de la ecuación del calor para un dominio acotado, a saber (*cf.* [10]),

$$u = \sum_{n=1}^{\infty} a_n e^{-kq_n^2 t} \cos q_n x, \tag{21}$$

donde las q_n están determinadas por las condiciones de frontera y las a_n por las condiciones iniciales. Fourier considera ahora las q_n como abscisas de una curva y las a_n como ordenadas de esa curva. Entonces $a_n = Q(q_n)$ donde Q es una función de q. Entonces reemplaza (21) por

$$u = \int_0^{\infty} Q(q) e^{-kq^2 t} \cos qx \, dq, \tag{22}$$

y busca determinar Q. Regresa a la fórmula para los coeficientes

$$a_n = \frac{2}{\pi} \int_0^{\pi} \phi(x) \cos nx \, dx,$$

donde ϕ (x) sería normalmente las función inicial. Por un «proceso de paso al límite», que reemplaza las a_n por Q y n por q, obtiene

$$Q = \frac{2}{\pi} \int_0^{\infty} F(x) \cos qx \, dx, \tag{23}$$

donde $F(x)$, una función par, es la temperatura dada inicial en el dominio infinito. Entonces, usando (23) en (22) y por un intercambio del orden de integración, que Fourier no se preocupa de cuestionar, obtiene

$$u = \frac{2}{\pi} \int_0^{\infty} F(\alpha) \, d\alpha \int_0^{\infty} e^{-kq^2 t} \cos qx \cos q\alpha \, dq.$$

Fourier, entonces, hace algo análogo para una $F(x)$ impar, y finalmente obtiene

$$u = \frac{1}{\pi} \int_{-\infty}^{\infty} F(\alpha) \, d\alpha \int_0^{\infty} e^{-kq^2 t} \cos q(x - \alpha) \, dq. \tag{24}$$

Así, la solución se expresa en forma cerrada. Ahora, para $t = 0$, u es $F(x)$, que puede ser cualquier función dada. De aquí Fourier asegura que, para una $F(x)$ arbitraria,

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} F(\alpha)\, d\alpha \int_{0}^{\infty} \cos q(x - \alpha)\, dq, \tag{25}$$

y ésta es una forma de representación de la integral doble de Fourier de una función arbitraria. En su libro, Fourier mostró cómo resolver muchos tipos de ecuaciones con su integral. Un uso se apoya en el hecho que, si (24) es obtenida por cualquier proceso, entonces (25) muestra que u satisface la condición inicial en $t = 0$. Otro empleo es más evidente si uno escribe la integral de Fourier en forma exponencial, usando la relación de Euler, $e^{ix} = \cos x + i \operatorname{sen} x$. Entonces (25) se transforma en

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iqx}\, dq \int_{-\infty}^{\infty} F(\alpha) e^{-iq\alpha}\, d\alpha.$$

Esta forma muestra que $F(x)$ puede ser descompuesta en un número infinito de componentes armónicos con una frecuencia $q/2\pi$ variando continuamente y con una amplitud $(1/2\pi) \int F(\alpha) e^{-iq\alpha}\, d\alpha$, mientras que la serie ordinaria de Fourier descompone una función dada en un conjunto infinito, pero discreto, de componentes armónicas.

La derivación de Cauchy de la integral de Fourier es de alguna manera similar. El artículo en el que apareció, «Théorie de la propagation des ondes» («Teoría de la propagación de ondas»), recibió el premio de la Academia de París de 1816 [5]. Este artículo es la primera investigación profunda sobre las ondas en la superficie de un fluido, una materia iniciada por Laplace en 1778. A pesar de que Cauchy establece las ecuaciones hidrodinámicas generales, se limita casi inmediatamente a casos especiales. En particular, considera la ecuación

$$\frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial y^2} = 0,$$

[5] *Mém. divers savans*, 1, 1827, 3-312 = *Œuvres* (1), 1, 5-318; véase, también, Cauchy, *Nouv. Bull. de la Soc. Phil.*, 1817, 121-124 = *Œuvres* (2), 2, 223-227.

donde q es lo que más tarde se llamó un potencial de velocidad, siendo x e y las coordenadas espaciales. Escribe la solución sin explicación alguna (*cf.* [22]).

$$q = \int_0^\infty \cos mx \, e^{-ym} f(m) \, dm, \tag{26}$$

donde $f(m)$ es arbitraria hasta entonces. Ya que $y = 0$ sobre la superficie, q se reduce a una $F(x)$ dada,

$$F(x) = \int_0^\infty \cos mx \, f(m) \, dm. \tag{27}$$

Entonces Cauchy muestra que

$$f(m) = \frac{2}{\pi} \int_0^\infty \cos mu \, F(u) \, du. \tag{28}$$

Con este valor de $f(m)$

$$F(x) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \cos mx \cos mu \, F(u) \, du \, dm. \tag{29}$$

Así, Cauchy obtiene no solamente la representación de la integral doble de Fourier de $F(x)$, sino que también da la transformada de Fourier de $f(m)$ a $F(x)$ y la transformada inversa. Dada $F(x)$, $f(m)$ esta determinada por (28) y puede ser usada en (26).

Poco tiempo después de que Cauchy entregara su ensayo ganador, Poisson, quien no podía competir por el premio, pues era miembro de la Academia, publicó un trabajo importante sobre ondas de agua, «Mèmoire sur la théorie des ondes» («Memoria sobre la teoría de ondas») [6]. En su trabajo deriva la integral de Fourier casi de la misma manera que Cauchy.

[6] *Mém. de l'Acad. des Sci., Paris* (2), 1, 1816, 71-186.

4. La ecuación del potencial y el teorema de Green

El siguiente avance significativo se centró en la ecuación del potencial, aunque el resultado principal, el teorema de Green, tiene aplicaciones en muchos otros tipos de ecuaciones diferenciales. La ecuación del potencial había figurado en el trabajo del siglo XVIII sobre gravitación y también en los estudios sobre conducción del calor del siglo XIX, referentes a que, cuando la distribución de la temperatura en un cuerpo, aunque varíe de punto a punto, permanece igual a pesar de que el tiempo varía, o cuando permanece en estado de reposo, entonces T en (1) es independiente del tiempo y la ecuación del calor se reduce a la ecuación del potencial. El interés puesto en la ecuación del potencial para el cálculo de la atracción gravitacional continuó al inicio del siglo XIX, pero acentuándose por una nueva clase de aplicaciones a la electrostática y magnetostática. Aquí, también, la atracción de los elipsoides constituyó un problema clave.

Una corrección en la teoría de la atracción gravitacional, expresada por la ecuación del potencial, fue hecha por Poisson [7]. Laplace (cap. 22, sec. 4) había supuesto que la ecuación del potencial

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0, \tag{30}$$

donde V es una función de x, y y z, se cumple en cualquier punto (x,y,z), ya sea dentro o fuera del cuerpo que ejerce la atracción gravitacional. Poisson demostró que si (x,y,z) cae dentro del cuerpo atrayente, entonces V satisface

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi\varrho, \tag{31}$$

donde ϱ es la densidad del cuerpo atrayente y es también función de x, y y z. A pesar de que (31) se llama aún ecuación de Poisson, la prueba de que se verifica no fue rigurosa, como reconoció, aún para los criterios de su tiempo.

En este mismo artículo Poisson llamó la atención sobre la utilidad de la función V en investigaciones eléctricas, señalando que su

[7] *Nouv. Bull. de la Soc. Philo.*, 3, 1813, 388-392.

valor sobre la superficie de cualquier conductor debe ser constante cuando la carga eléctrica está distribuida sobre toda la superficie. En otros artículos resolvió un buen número de problemas relativos a la distribución de la carga sobre las superficies de cuerpos conductores cuando los cuerpos se encuentran cerca uno de otro. Su principio básico fue que la fuerza electrostática resultante en el interior de cualquiera de los conductores debía ser cero.

A pesar del trabajo de Laplace, Poisson, Gauss y otros, casi nada se conocía, hacia 1820, acerca de las propiedades generales de las soluciones de la ecuación del potencial. Se pensaba que la integral general debía contener dos funciones arbitrarias, de las cuales una proporciona el valor de la solución sobre la frontera, y la otra, la derivada de la solución sobre la frontera. Sin embargo, se conocía, en el caso de la conducción de calor en estado estable, en la que la temperatura satisface la ecuación del potencial, que la temperatura o distribución del calor a través de un cuerpo tridimensional está determinada cuando sólo la temperatura está determinada sobre la superficie. De aquí que una de las funciones arbitrarias en la supuesta solución general de la ecuación del potencial debe estar dada por alguna otra condición.

En este punto, George Green (1793-1841), un matemático inglés autodidacta, se dedicó a tratar la electricidad estática y el magnetismo de una manera completamente matemática. En 1828, Green publicó un panfleto impreso en forma privada, *An essay on the application of Mathematical Analysis to the Theories of Electricity and Magnetism (Un ensayo sobre la aplicación del análisis matemático a las teorías de la electricidad y el magnetismo)*. Esto fue ignorado hasta que sir William Thomson (lord Kelvin, 1824-1907) lo descubrió, reconoció su gran valor y lo publicó en el *Journal für Mathematic* [8]. Green, quien aprendió de los artículos de Poisson, llevó también la noción de función potencial a la electricidad y al magnetismo.

Empezó con (30) y demostró los siguientes teoremas: Sean U y V dos funciones continuas cualesquiera de x, y y z cuyas derivadas no son infinitas en ningún punto de un cuerpo arbitrario. El principal teorema asegura que (usaremos ΔV para el lado izquierdo de (30), aunque no fue usado por Green)

---

[8] *Jour. für Math.*, 39, 1850, 73-89; 44, 1852, 356-374; y 47, 1854, 161-221 = En los *Mathematical Papers* de Green, 1871, 3-115.

$$\iiint U\,\Delta V\,dv + \iint U\frac{\partial V}{\partial n}\,d\sigma$$

$$= \iiint V\,\Delta U\,dv + \iint V\frac{\partial U}{\partial n}\,d\sigma, \qquad (32)$$

donde n es la normal a la superficie del cuerpo dirigida hacia adentro y $d\sigma$ un elemento de superficie. El teorema (32), incidentalmente, también lo demostró Miguel Ostrogradsky (1801-1861), un matemático ruso, quien los presentó a la Academia de Ciencias de San Petesburgo en 1828 [9].

Green demostró más adelante que el requisito de que V y cada una de sus primeras derivadas fueran continuas en el interior del cuerpo puede ser impuesto en lugar de una condición de frontera sobre las derivadas de V. A la luz de este hecho, Green representó V en el interior del cuerpo, en términos de su valor $\bar{V}$ sobre la frontera (función que estaría dada) y en términos de otra función U que tiene las propiedades: (a) U debe ser 0 sobre la superficie; (b) en un punto P fijo pero indeterminado en el interior, U se hace infinita como $1/r$, donde r es la distancia de cualquier otro punto a partir de P; (c) U debe satisfacer la ecuación del potencial (30) en el interior. Si U es conocida, y podría ser más fácilmente encontrada porque satisface condiciones más simples que V, entonces V puede ser representada en todo punto interior por

$$4\pi V = -\iint \bar{V}\frac{\partial U}{\partial n}\,d\sigma,$$

donde la integral se extiende sobre la superficie, y $\partial U/\partial n$ es la derivada de U en la dirección perpendicular a la superficie hacia el cuerpo. Se entiende que las coordenadas de P están contenidas en $\partial U/\partial n$ y son los argumentos en P. Esta función U, introducida por Green, que Riemann llamó más tarde función de Green, se convirtió en un concepto fundamental de las ecuaciones en derivadas parciales. El propio Green usó el término «función potencial» para esta función especial U, tanto como para V. Su método de obtener soluciones de la ecuación del potencial, como contrario al método de usar series de funciones especiales, es llamado el método de las singula-

[9] *Mém. Acad. Sci. St. Peters.* (6), 1, 1831, 39-53.

ridades. Desgraciadamente, no existe una expresión general para la función U, como tampoco existe un método para encontrarla. Green se contentaba en esta materia con proporcionar el significado físico de U para el caso del potencial creado por cargas eléctricas.

Green aplicó este teorema y sus conceptos a problemas eléctricos y magnéticos. En 1833 [10] estudió también el problema del potencial gravitacional para elipsoides de densidades variables. En su trabajo, Green demostró que cuando V está dada sobre la frontera de un cuerpo, existe sólo una función que satisface $\Delta V = 0$ en todo el cuerpo, no tiene singularidades y toma los valores dados en la frontera. Para hacer esta prueba, Green asumió la existencia de una función que minimiza

$$\iiint \left[\left(\frac{\partial V}{\partial x} \right)^2 + \left(\frac{\partial V}{\partial y} \right)^2 + \left(\frac{\partial V}{\partial z} \right)^2 \right] dv. \tag{33}$$

Este es el primer uso del principio de Dirichlet (*cf.* cap. 27, sec. 8).

En su ensayo de 1835, Green realizó gran parte del trabajo en n dimensiones en lugar de tres y también proporcionó importantes resultados sobre lo que nosotros llamamos ahora funciones ultraesféricas, que son generalizaciones de n variables de los armónicos esféricos de Laplace. Ya que la obra de Green no fue bien conocida por mucho tiempo, algunos otros hicieron parte de su trabajo independientemente.

Green es el primero de los grandes matemáticos británicos en seguir las corrientes del trabajo hecho en el continente después de la introducción del análisis en Inglaterra. Su trabajo inspiró la gran escuela de físico-matemáticos de Cambridge, que incluía a sir William Thompson, sir Gabriel Stokes, lord Rayleigh y Clerk Maxwell.

Los logros de Green fueron seguidos por la obra maestra de Gauss de 1839 [11], «Allgemeine Lehrsätze in Beziehung auf die im verkehrten Verhaltnisse des Quadrats der Entfernung wirkenden Anziehungs-und Abstossungs-krafte»» («Teoremas generales sobre fuerzas atractivas y repulsivas que actúan de acuerdo con el inverso del cuadrado de la distancia»). Gauss demostró rigurosamente el resultado de Poisson, a saber, que $\Delta V = -4\pi\varrho$ en un punto dentro

[10] *Trans. Camb. Phil. Soc.*, 5, 1835, 395-430 = *Mathematical Papers*, 187-222.

[11] *Resultate aus den Beobachtungen des magnestischen Vereins*, vol. 4, 1840 = *Werke*, 5, 197-242.

de la masa actuante, bajo la condición de que ρ es continua en ese punto y en un pequeño dominio alrededor de él. Esta condición no se satisface sobre la superficie de la masa actuante. En la superficie las cantidades $\partial V/\partial x^2$, $\partial^2 V/\partial y^2$, y $\partial^2 V/\partial z^2$ tienen saltos.

Hasta ahora el trabajo realizado en torno a la ecuación del potencial y la ecuación de Poisson suponía la existencia de una solución. La prueba de Green de la existencia de una función de Green se apoyaba por completo en un argumento físico. Desde el punto de vista de la existencia, el problema fundamental de la teoría del potencial era mostrar la existencia de una función potencial V, la que William Thomson, alrededor de 1850, llamó una función armónica, cuyos valores están dados sobre la frontera de una región y que satisface $\Delta V = 0$ en la región. Se podía establecer esto directamente, o establecer la existencia de la función U de Green y de ahí obtener V. El problema de establecer la existencia de la función de Green o de la misma V es conocido como el problema de Dirichlet o el primer problema de contorno de la teoría del potencial, el problema de existencia más básico y antiguo de la materia. El problema de encontrar una V que satisface $\Delta V = 0$ en una región cuando la derivada normal de V es especificada sobre la frontera, es llamado problema de Neumann, en homenaje a Carl G. Neumann (1832-1925), un profesor de Leipzig. A este problema se le denomina segundo problema fundamental de la teoría del potencial.

Una aproximación al problema de establecer la existencia de una solución para $\Delta V = 0$, que Green ya había utilizado (véase [33]), fue puesta en evidencia por William Thomson. En 1847 [12], Thomson anunció el teorema o principio que en Inglaterra lleva su nombre y que en el continente se menciona como el *principio* de Dirichlet, ya que así lo llamó Riemann. A pesar de que Thomson lo enunció en una forma un poco más general, la esencia del principio puede ser expuesta así: considérese la clase de todas las funciones U que tienen derivadas de segundo orden continuas en los dominios interior y exterior T y T', respectivamente, separados por una superficie S. Las U deben ser continuas en todas partes y tomar sobre S los valores de una función continua f. La función V que minimiza la integral de Dirichlet

---

12 *Jour. de Math.*, 12, 1847, 493-496 = *Cambridge and Dublin Math. Jour.*, 3, 1848, 84-87 = *Math. and Physical Papers*, 1, 93-96.

$$I = \iiint_T \left[\left(\frac{\partial U}{\partial x} \right)^2 + \left(\frac{\partial U}{\partial y} \right)^2 + \left(\frac{\partial U}{\partial z} \right)^2 \right] dv \qquad (34)$$

es la que satisface $\Delta V = 0$ y toma el valor f sobre la frontera S. La conexión entre (34) y ΔV es que la primera variación de I en el sentido del cálculo de variaciones es ΔV, y ésta debe ser 0 para una V minimizante. Ya que para U real la I no puede ser negativa, parecía claro que habría de existir una función minimizante V, y entonces no es difícil probar que es única. El principio de Dirichlet es entonces *una* forma de aproximarse al problema de Dirichlet de la teoría del potencial.

El trabajo de Riemann sobre funciones complejas dio una nueva importancia al problema de Dirichlet y al principio en sí mismo. La «prueba» de Riemann de la existencia de V en su tesis doctoral usó el caso bidimensional del principio de Dirichlet, pero no era rigurosa, de lo que se dio cuenta él mismo.

Cuando en su ensayo de 1870 [13], Weierstrass presentó una crítica del principio de Dirichlet, mostró que la existencia *a priori* de una U minimizante no estaba apoyada por argumentos sólidos. Era correcto que para todas las funciones U diferenciablemente continuas que van continuamente del interior a los valores de frontera prescritos, la integral tiene una *cota inferior*. Pero que exista una función U_0 en la clase de funciones diferenciables continuas que alcance la cota inferior, eso no fue establecido.

Otra técnica para la solución de la ecuación del potencial emplea la teoría de funciones complejas. A pesar de que D'Alembert, en su trabajo de 1752 (cap. 27, sec. 2), y Euler en problemas especiales, habían usado esta técnica para resolver la ecuación del potencial, no fue sino hasta mediados del siglo XIX cuando la teoría de funciones complejas se empleó vitalmente en la teoría del potencial. La relevancia de la teoría de funciones para la teoría del potencial se apoya en el hecho de que si $u + iv$ es una función analítica de z, entonces ambas u y v satisfacen la ecuación de Laplace. Más aún, si u satisface la ecuación de Laplace, entonces la función conjugada v tal que $u + iv$ es analítica necesariamente existe (cap. 27, sec. 8).

Cuando la ecuación $\Delta u = 0$ se utiliza en el flujo de fluidos, la función $u(x,y)$ es lo que Helmholtz llamó potencial de velocidades

[13] Cap. 27, sec. 9.

y entonces $\partial u/\partial x$ y $\partial u/\partial y$ representan las componentes de la velocidad del fluido en cualquier punto (x,y). En el caso de la electricidad estática, u es el potencial electrostático y $\partial u/\partial x$ y $\partial u/\partial y$ son las componentes de la fuerza eléctrica. En ambos casos las curvas u = const. son líneas equipotenciales y las curvas v = const., que son ortogonales a u = const., son el flujo o líneas de corriente (líneas de fuerza para la electricidad). A la función $v(x,y)$ se la llama función de corriente. La introducción de esta función, claramente, es una ayuda por su significado físico.

Una de las ventajas del uso de la teoría de las funciones complejas en la solución de la ecuación del potencial se deriva del hecho de que si $F(z) = F(x + i\,y)$ es una función analítica, de tal manera que sus partes real e imaginaria satisfacen $\Delta V = 0$, entonces la transformación de x e y en ξ y η por

$$\xi = f(x, y), \qquad \eta = g(x, y)$$

donde

$$\zeta = \xi + i\eta$$

nos da otra función analítica $G(\zeta) = G(\xi + i\eta)$, y sus partes real e imaginaria también satisfacen $\Delta V(\xi,\eta) = 0$. Ahora bien, si el problema original del potencial $\Delta V = 0$ tiene que ser resuelto en algún dominio D, entonces por la adecuada elección de la transformación, el dominio D', en el que la transformada $\Delta V = 0$ tiene que ser resuelta puede ser mucho más simple. Aquí el uso de transformaciones conformes tal como la transformación de Schwarz-Christoffel, es de gran utilidad.

No vamos a seguir el uso de la teoría de las funciones complejas en la teoría del potencial porque los detalles de su uso van más allá de cualquier metodología básica en la solución de ecuaciones diferenciales parciales. Sin embargo, de nuevo es importante mencionar que muchos matemáticos se resistieron al uso de funciones complejas porque todavía no se reconciliaban con los números complejos. En la universidad de Cambridge, aún en 1850, se utilizaban artificios engorrosos a fin de evitar el uso de las funciones complejas. El libro de Horace Lamb *Treatise on the Mathematical Theory of the Motion of Fluids (Tratado sobre la teoría matemática del movimiento de fluidos)*, publicado en 1879 y todavía un clásico (ahora conocido

como *Hydrodynamics (Hidrodinámica)*), fue el primer texto en reconocer la aceptación de la teoría de funciones en Cambridge.

5. Coordenadas curvilíneas

Green introdujo un número considerable de ideas importantes cuya significación se extendía más allá de la ecuación del potencial. Gabriel Lamé (1795-1870), matemático e ingeniero preocupado en principio por la ecuación del calor, introdujo otra técnica importante: se valió de un sistema de coordenadas curvilíneas, que también podía ser usado para muchos tipos de ecuaciones. Lamé señaló en 1833 [14] que la ecuación del calor había sido resuelta únicamente para cuerpos conductores cuyas superficies son normales a los planos coordenados x = const., y = const. y z = const. La idea de Lamé consistió en introducir nuevos sistemas de coordenadas y las correspondientes coordenadas en la superficie. Con ciertas restricciones esto lo habían hecho Euler y Laplace, quienes usaron coordenadas esféricas ϱ, θ y ϕ, en cuyo caso las superficies coordenadas ϱ = const., θ = const. y ϕ = const. son esferas, planos y conos, respectivamente. Conociendo las ecuaciones que transforman las coordenadas rectangulares en esféricas es posible, como lo fue para Euler y Laplace, tranformar la ecuación del potencial de coordenadas rectangulares a esféricas.

El valor de los nuevos sistemas de coordenadas y superficies tiene dos sentidos. Primero, una ecuación diferencial parcial en coordenadas rectangulares puede no ser separable en ecuaciones diferenciales ordinarias en este sistema, pero podría ser separable en algún otro sistema. Segundo, el problema físico podría requerir una condición de contorno, digamos, sobre una elipsoide. Tal frontera, simplemente es representada en un sistema coordenado donde una familia de superficies consiste de elipsoides, mientras que en un sistema rectangular se usa una ecuación relativamente complicada. Además, después que es empleada la separación de variables en el propio sistema de coordenadas, esta condición de frontera se aplica a sólo una de las ecuaciones diferenciales ordinarias resultantes.

Lamé introdujo diversos nuevos sistemas de coordenadas con el propósito expreso de resolver la ecuación de calor en estos siste-

[14] *Jour de l'Ecole Poly.*, 14, 1833, 194-251.

mas [15]. Su sistema principal fueron las tres familias de superficies dadas por las ecuaciones

$$\begin{aligned}
&\frac{x^2}{\lambda^2} + \frac{y^2}{\lambda^2 - b^2} + \frac{z^2}{\lambda^2 - c^2} - 1 = 0 \\
&\frac{x^2}{\mu^2} + \frac{y^2}{\mu^2 - b^2} + \frac{z^2}{\mu^2 - c^2} - 1 = 0 \qquad (35) \\
&\frac{x^2}{\nu^2} + \frac{y^2}{\nu^2 - b^2} + \frac{z^2}{\nu^2 - c^2} - 1 = 0,
\end{aligned}$$

donde $\lambda^2 > c^2 > \mu^2 > b^2 > \nu^2$. Estas tres familias son elipsoides, hiperboloides de una hoja e hierpoboloides de dos hojas, todas las cuales tienen el mismo foco. Cualquier superficie de una familia corta todas las superficies de las otras dos ortogonalmente, y de hecho las corta en las líneas de curvatura (cap. 23, sec. 7). Así, cualquier punto en el espacio tiene coordenadas (λ, μ, ν), a saber, las λ, μ, y ν de las superficies, una de cada familia, que pasan por ese punto. Este nuevo sistema de coordenadas es llamado elipsoidal, aunque Lamé lo denominó eliptical, término usado hoy día para otro sistema.

Lamé transformó la ecuación de calor para el caso del estado estacionario (la temperatura es independiente del tiempo), esto es, la ecuación del potencial, a estas coordenadas, y demostró que podía usar la separación de variables para reducir la ecuación a tres ecuaciones diferenciales ordinarias. Por supuesto, estas ecuaciones deben ser resueltas sujetas a las apropiadas condiciones en la frontera. En un ensayo de 1839 [16], Lamé estudió aún más la distribución de la temperatura en el estado estacionario en un elipsoide de tres ejes, y dio una solución completa del problema tratado en su artículo de 1833. En este artículo de 1839 introdujo también otro sistema de coordenadas curvilíneas, ahora llamado el sistema esferoconal, donde las superficies coordenadas son una familia de esferas y dos familias de conos. Lamé utilizó también este sistema para resolver problemas de conducción de calor. Lamé escribió muchos otros ensayos sobre la conducción del calor utilizando coordenadas elipsoidales, inclu-

---

15 *Annales de Chimie et Physique* (2), 53, 1833, 190-204.
16 *Jour. de Math.*, 4, 1839, 126-163.

yendo un segundo de 1839 en el mismo volumen del *Journal de Mathématiques*, donde trata casos especiales del elipsoide [17].

La cuestión de las familias de superficies mutuamente ortogonales tenía tan obvia importancia en la solución de ecuaciones en derivadas parciales que se convirtió en una materia de investigación en sí y por sí misma. En un artículo de 1834 [18] Lamé consideró las propiedades generales de tres familias cualesquiera de superficies mutuamente ortogonales y proporcionó un procedimiento para expresar una ecuación diferencial parcial en cualquier sistema ortogonal de coordenadas, una técnica que se ha venido utilizando continuamente desde entonces.

(Heinrich) Eduard Heine (1821-1881) siguió las huellas de Lamé. Heine, en su tesis doctoral de 1842 [19], determinó el potencial (temperatura en el estado estacionario) no únicamente para el interior de una elipsoide de revolución cuando el valor del potencial está dado sobre la superficie, sino también para el exterior de dicho elipsoide y para la concha entre elipsoides cofocales de revolución.

Lamé estaba tan impresionado por lo que él y otros habían logrado mediante el uso de sistemas de coordenadas triplemente ortogonales que pensó que todas las ecuaciones diferenciales parciales se podían resolver encontrando un sistema adecuado. Más tarde se dio cuenta de que esto era un error. En 1859 publicó un libro sobre la materia *Leçons sur les coordonnées curvilignes (Lecciones sobre las coordenadas curvilíneas)*.

Aunque del uso de familias mutuamente ortogonales de superficies como superficies coordenadas no resolvió todas las ecuaciones diferenciales parciales, abrió una nueva técnica que podía explotarse ventajosamente en muchos problemas. El uso de coordenadas curvilíneas se llevó a otras ecuaciones diferenciales parciales. Así, Emile Leonard Mathieu (1835-1900), en un ensayo de 1868 [20], trató las vibraciones de una membrana elíptica, que involucra la ecuación de ondas, e introdujo coordenadas cilíndricas elípticas y funciones apropiadas para estas coordenadas, ahora llamadas funciones de Mathieu (cap. 29, sec. 2). En el mismo año, Heinrich Weber (1842-1913), trabajando en la ecuación $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + k^2 u = 0$, la resolvió [21]

[17] *Jour. de Math.*, 4, 1839, 351-385.
[18] *Jour. de l'Ecole Poly.*, 14, 1834, 191-288.
[19] *Jour. für Math.*, 26, 1843, 185-216.
[20] *Jour. de Math.* (2), 13, 1868, 137-203.
[21] *Math. Ann.*, 1, 1869, 1-36.

para un dominio limitado por una elipse completa y también para la región limitada por los dos arcos de elipses cofocales y dos arcos de hipérbolas cofocales con las elipses. El caso especial en el que las elipses se convierten en parábolas cofocales fue tratado también, y aquí Weber introdujo funciones apropiadas para las expansiones en este sistema de coordenadas, ahora llamadas funciones de Weber o funciones cilíndricas parabólicas. En su *Cours de physique mathématique* (*Curso de física matemática,* 1873), Mathieu atacó nuevos problemas usando elipsoides e introdujo otras funciones nuevas.

Nuestra discusión —iniciada con Lamé— en torno al uso de coordenadas curvilíneas, describe únicamente el inicio de su trabajo. Muchos otros sistemas de coordenadas se habían introducido ya, al igual que funciones especiales resultantes de resolver ecuaciones diferenciales ordinarias que surgen de la separación de variables [22]. La mayor parte de esta teoría de funciones especiales fue creada por físicos, tal y como se iban necesitando de acuerdo a sus propiedades en problemas concretos (véase también el cap. 29).

6. La ecuación de ondas y la ecuación de ondas reducida

Tal vez el tipo más importante de ecuación diferencial parcial es la ecuación de ondas. En tres dimensiones espaciales la forma básica es

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = a^2 \frac{\partial^2 u}{\partial t^2} . \tag{36}$$

Como sabemos, esta ecuación ya había sido introducida en el siglo XVIII y también se expresó en coordenadas esféricas. Durante el siglo XIX se encontraron nuevas aplicaciones de la ecuación de ondas, especialmente en el campo vigoroso de la elasticidad. Las vibraciones de cuerpos sólidos con diversas formas de contorno, con diferentes condiciones iniciales y de frontera, y la propagación de ondas en cuerpos elásticos produjeron multitud de problemas. Trabajo subsiguiente en la propagación del sonido y la luz originó cientos de problemas adicionales.

---

[22] Véase: William E., Byerly, *An Elementary Treatise on Fourier Series,* Dover (reim.), 1959, y E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics,* Chelsea (reim.), 1955.

Donde la separación de variables es posible, la técnica para resolver (36) no es diferente de lo que hizo Fourier con la ecuación del calor o lo hecho por Lamé después de expresar la ecuación del potencial en algún sistema de coordenadas curvilíneas. El uso por Mathieu de las coordenadas curvilíneas para resolver la ecuación de ondas mediante la separación de variables es típico de cientos de artículos.

Tratando la ecuación como un todo se obtuvo una clase distinta e importante de resultados para la ecuación de ondas. El primero de estos resultados importantes es para problemas de valores iniciales y se remonta hasta Poisson, quien trabajó sobre esta ecuación durante los años de 1808 a 1819. Su logro principal [23] fue la fórmula para la propagación de una onda $u(x,y,z,t)$ cuyo estado inicial está descrito por las condiciones iniciales

$$u(x, y, z, 0) = \phi_0(x, y, z), \qquad u_t(x, y, z, 0) = \phi_1(x, y, z) \tag{37}$$

y que satisface la ecuación diferencial parcial

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{a^2}\frac{\partial^2 u}{\partial t^2} \tag{38}$$

donde a es una constante. La solución u está dada mediante

$$\begin{aligned} u(x, y, z, t) & \\ = \frac{1}{4\pi a}\int_0^{\pi}\int_0^{2\pi} & \phi_1(x + at \operatorname{sen}\phi\cos\theta, y + at \operatorname{sen}\phi \operatorname{sen}\theta, \\ & \qquad z + at\cos\phi)at \operatorname{sen}\phi \, d\theta \, d\phi \\ + \frac{1}{4\pi a}\frac{\partial}{\partial t}\int_0^{\pi}\int_0^{2\pi} & \phi_0(x + at \operatorname{sen}\phi\cos\theta, y + at \operatorname{sen}\phi \operatorname{sen}\theta, \\ & \qquad z + at\cos\phi)at \operatorname{sen}\phi \, d\theta \, d\phi \end{aligned} \tag{39}$$

donde θ y ϕ son las coordenadas esféricas usuales. El dominio de integración es la superficie de la esfera S_{at} con radio at alrededor del punto P de coordenadas x, y y z.

---

[23] *Mém. de l'Acad. des Sci., Paris* (2), 3, 1818, 121-176.

Para obtener alguna indicación de lo que significa el resultado de Poisson, consideremos un ejemplo físico. Supongamos que la perturbación inicial está originada por un cuerpo V (Fig. 28.2) con frontera S de tal manera que ϕ_0 y ϕ_1 están definidas sobre V y son 0 fuera de V. Nosotros decimos que la perturbación original está localizada en V. Físicamente, una onda sale de V y se extiende en el espacio. La fórmula de Poisson nos dice lo que sucede en cualquier punto $P(x,y,z)$ fuera de V. Sean d y D las distancias mínima y máxima de P a los puntos de V. Cuando $t < d/a$, las integrales en (39) son 0 ya que el dominio de integración es la superficie de la esfera S_{at} con radio at y centro en P. Puesto que ϕ_0 y ϕ_1 son 0 en S_{at}, entonces la función u es 0 en P. Esto significa que la onda que sale de S no ha llegado a P. En $t = d/a$, la esfera S_{at} únicamente toca a S, así que el frente de la onda emanando de S llega precisamente a P. Entre $t = d/a$ y $t = D/a$ la esfera S_{at} corta a V y de ese modo $u(P,t) \neq 0$. Finalmente, para $t > D/a$, la esfera S_{at} no cortará a S (la región entera V cae dentro de S_{at}); es decir, la perturbación inicial ha pasado a través de P. Así, de nuevo $u(P,t) = 0$. El instante $t = D/a$ corresponde al paso del borde de la onda a través de P. En cualquier tiempo dado t, el borde anterior de la onda toma la forma de una superficie que separa los puntos a los que ha llegado la perturbación de aquellos a los que no lo ha hecho. Este borde anterior es la envolvente de la familia de esferas con centros sobre S y con radios at. El borde posterior de la onda en el tiempo t es una superficie que separa puntos en los que la perturbación existe de aquellos en los que la perturbación ya pasó. Vemos, entonces, que la perturbación que está localizada en el espacio da origen en cada punto P a un efecto que sólo dura por un tiempo finito. Más aún, la (pertur-

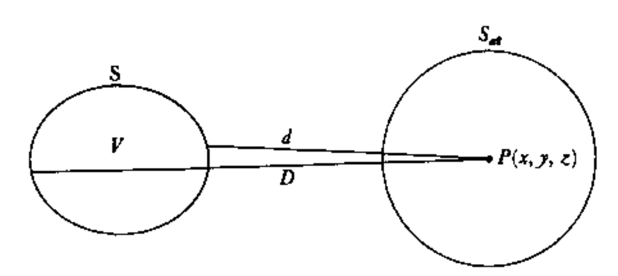


FIGURA 28.2

bación) de onda tiene un borde anterior y uno posterior. Este fenómeno, en su totalidad, es llamado el principio de Huygens.

Una forma bastante diferente de resolver problemas de valor inicial para la ecuación de ondas fue creado por Riemann en el transcurso de su trabajo sobre la propagación de ondas de sonido de amplitud finita [24]. Riemann considera una ecuación diferencial lineal de segundo orden que puede ser puesta en la forma

$$L(u) = \frac{\partial^2 u}{\partial x\,\partial y} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu = 0, \qquad (40)$$

donde D, E y F son funciones de x e y continuas y diferenciables hasta el orden dos. El problema requiere encontrar u en un punto arbitrario P (Fig. 28.3) cuando se conoce u y $\partial u/\partial n$ (lo que significa conocer $\partial u/\partial x$ y $\partial u/\partial y$ a lo largo de la curva Γ). Su método depende de encontrar una función v (llamada función de Riemann o función característica) [25] que satisface lo que ahora se llama la ecuación adjunta

$$M(v) = \frac{\partial^2 v}{\partial x\,\partial y} - \frac{\partial(Dv)}{\partial x} - \frac{\partial(Ev)}{\partial y} + Fv = 0 \qquad (41)$$

y otras condiciones que especificaremos en breve.

Riemann introdujo los segmentos PP_1 y PP_2 de las características (él no usó este término) $x = \xi$, $y = \eta$ pasando por P. Ahora se aplica

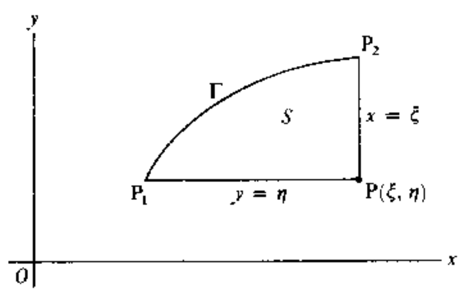


FIGURA 28.3

[24] *Abh. der Ges. der Wiss. zu Gött.*, 8, 1858-1859, 43-65 = *Werke*, 156-178.
[25] v no es la misma como solución fundamental o como función de Green.

una generalización del teorema de Green (en dos dimensiones) a la expresión diferencial $L(u)$. Para expresar el teorema en una forma resumida, introduzcamos

$$X = \frac{1}{2}\left(v\frac{\partial u}{\partial y} - u\frac{\partial v}{\partial y}\right) + Duv$$

$$Y = \frac{1}{2}\left(v\frac{\partial u}{\partial x} - u\frac{\partial v}{\partial x}\right) + Euv.$$

Entonces, el teorema de Green afirma que

$$\begin{aligned}\int_S [vL(u) - uM(v)]\, dS &= \int_S \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dS \\ &= \int_C \{X\cos(n, x) + Y\cos(n, y)\}\, ds,\end{aligned} \tag{42}$$

donde S es el área en la figura, C es la frontera entera de S y *cos* (n,x) es el coseno del angulo entre la normal a C y el eje x.

Además de satisfacer (41), Riemann requiere de v que cumpla

$$\begin{aligned}&\text{(a)}\quad v = 1 \quad \text{at} \quad P,\\ &\text{(b)}\quad \frac{\partial v}{\partial y} - Dv = 0 \quad \text{en} \quad x = \xi,\\ &\text{(c)}\quad \frac{\partial v}{\partial x} - Ev = 0 \quad \text{en} \quad y = \eta.\end{aligned} \tag{43}$$ [26]

Usando la condición de que $M(v) = 0$ y las condiciones (43), y evaluando la integral curvilínea sobre C, Riemann obtiene

$$u(\xi, \eta) = \int_\Gamma \{X\cos(n, x) + Y\cos(n, y)\}ds + \frac{1}{2}\{(uv)_{P_1} + (uv)_{P_2}\}. \tag{44}$$

Así el valor de u en cualquier punto arbitrario P está dado en términos de los valores de u, $\partial u/\partial n$, v, y $\partial v/\partial n$ sobre Γ y los valores de u y v en P_1 y P_2.

---

[26] Para problemas bidimensionales v es una función de cuatro variables, ξ, η, x, e y. Satisface $M(v) = 0$ como función de x e y.

Ahora, u está dada en P_1 y P_2. La función v debe ser encontrada resolviendo $M(v) = 0$ y cumpliendo las condiciones en (43). De esta forma, lo que logra el método de Riemann es cambiar el problema original del valor inicial para u en otro tipo de problema de valor inicial, ahora para v. Generalmente, el segundo problema es más fácil de resolver. En el problema físico de Riemann fue especialmente fácil encontrar v. Sin embargo, la existencia general de tal v no fue establecida por Riemann.

El método de Riemann, como fue descrito con anterioridad, es útil únicamente para el tipo de ecuación especificada por la ecuación de ondas (ecuaciones hiperbólicas) en dos variables independientes y no podía ser extendido directamente. La extensión del método a más de dos variables independientes se enfrenta a la dificultad de que la función de Riemann se hace singular sobre la frontera del dominio de integración y la integral diverge. El método ha sido extendido, pero incrementando las complicaciones.

El progreso en la solución de la ecuación de ondas por otros métodos está íntimamente conectado con los llamados problemas de estado estacionario que condujeron a la ecuación de onda reducida. La ecuación de onda, por su misma forma, incluye la variable tiempo. En muchos problemas físicos, donde uno está interesado en ondas armónicas simples, se supone que $u = w(x,y,z)e^{ikt}$, y sustituyendo esto en la ecuación de ondas se obtiene

$$\Delta w + k^2 w = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} + k^2 w = 0. \qquad (45)$$

Esta es la ecuación de ondas reducida o ecuación de Helmholtz. La ecuación $\Delta w + k^2 w = 0$ representa todas las ondas armónicas, elásticas acústicas y electromagnéticas. Mientras que a los autores anteriores les bastaba encontrar integrales particulares, Hermann von Helmholtz (1821-1894), en su trabajo sobre las oscilaciones del aire en un tubo (de órgano), con un extremo abierto, proporcionó la primera investigación general de sus soluciones [27]. Estaba interesado por el problema acústico en el cual w es el potencial de velocidades de masas de aire moviéndose armónicamente, k es una constante determinada por la elasticidad del aire y la frecuencia de oscilación y λ, que es igual a $2\pi/k$, es la longitud de onda. Aplicando el teo-

[27] *Jour. für Math.*, 57, 1860, 1-72 = *Wissenschaftliche Abhandlungen*, 1, 303-382.

rema de Green, demostró que cualquier solución de $\Delta w + k^2 w = 0$ que es continua en un dominio dado puede ser representada como el efecto de capas sencillas o dobles de puntos de excitación sobre la superficie del dominio. Usando $e^{-ikr}/4\pi r$ como una de las funciones del teorema de Green, obtuvo

$$w(P) = -\frac{1}{4\pi}\iint \frac{e^{-ikr}}{r}\frac{\partial w}{\partial n}\,dS + \frac{1}{4\pi}\iint w\,\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)dS \qquad (46)$$

donde r denota la distancia desde P a un punto variable de la frontera. Así w está dado en cualquier punto P dentro del dominio en el que la solución es buscada en términos de los valores de w y $\partial w/\partial n$ sobre la frontera S.

Gustav R. Kirchhoff (1824-1887), uno de los más grandes físico-matemáticos alemanes del siglo XIX, se valió del trabajo de Helmholtz para obtener otra solución del problema de valor inicial para la ecuación de ondas. Supongamos que $\Delta w + k^2 w = 0$ viene de

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u$$

donde hemos puesto $u = we^{iat}$ de tal forma que $k = \sigma/c$. Entonces (46) puede ser escrita como

$$u(P, t) = -\frac{1}{4\pi}\iint \frac{e^{i\sigma[t-(r/c)]}}{r}\frac{\partial u}{\partial n}\,dS + \frac{1}{4\pi}\iint u\,\frac{\partial}{\partial n}\left(\frac{e^{i\sigma[t-(r/c)]}}{r}\right)dS. \qquad (47)$$

Esta fórmula fue generalizada por Kirchhoff. Si hacemos que $\phi(t)$ sea el valor de u en cualquier punto (x,y,z) de la frontera en el instante t y que $f(t)$ sea el valor correspondiente de $\partial u/\partial n$, entonces Kirchhoff demostró [28] que

$$u(P, t) = -\frac{1}{4\pi}\iint \frac{f[t-(r/c)]}{r}\,dS + \frac{1}{4\pi}\iint \frac{\partial}{\partial n}\left(\frac{\phi[t-(r/c)]}{r}\right)dS, \qquad (48)$$

siempre que en el último término la diferenciación con respecto a n se aplique a r solamente cuando aparece explícitamente en el nume-

[28] *Sitzungsber. Akad. Wiss. zu Berlin*, 1882, 641-669 = *Ges. Abh.*, 2, 22 sigs.

rador y el denominador. Así, u es obtenida en P en términos de los valores de u y $\partial u/\partial n$ en tiempos anteriores en puntos de la superficie cerrada rodeando P. A este resultado se llama el principio de Huygens de la acústica y representa una generalizacion de la fórmula de Poisson.

Hemos hecho notar que Riemann usó una especie de generalización del teorema de Green. La generalización ulterior del teorema de Green que emplea la ecuación diferencial adjunta y que es llamado también teorema de Green, proviene de un ensayo de Paul Du Bois Reymond (1831-1889)[29] y de la *Théorie générale des surfaces (Teoría general de superficies)* de Darboux[30]; ambos citan el ensayo de Riemann de 1858/1859. Si la ecuación dada es

$$L(u)=A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\,\partial y}+C\frac{\partial^2 u}{\partial y^2}+D\frac{\partial u}{\partial x}+E\frac{\partial u}{\partial y}+Fu=0,$$

donde los coeficientes son funciones de x e y, se integra el producto $vL(u)$ sobre un dominio arbitrario R del plano x, y bajo la suposición de que u, v y sus primeras y segundas derivadas parciales son continuas. Entonces la integración por partes lleva al teorema de Green generalizado, el cual establece que

$$\iint uM(v)\,dx\,dy=-\iint vL(u)\,dx\,dy-\int(Q\,dy-P\,dx),$$

donde las integrales dobles son calculadas sobre el interior de R y las integrales simples sobre la frontera de R,

$$M(v)=\frac{\partial^2(Av)}{\partial x^2}+2\frac{\partial^2(Bv)}{\partial x\,\partial y}+\frac{\partial^2(Cv)}{\partial y^2}-\frac{\partial(Dv)}{\partial x}-\frac{\partial(Ev)}{\partial y}+Fu$$

$$P=B\left(v\frac{\partial u}{\partial x}-u\frac{\partial v}{\partial x}\right)+C\left(v\frac{\partial u}{\partial y}-u\frac{\partial v}{\partial y}\right)+\left(E-\frac{\partial B}{\partial x}-\frac{\partial C}{\partial y}\right)uv$$

$$Q=A\left(v\frac{\partial u}{\partial x}-u\frac{\partial v}{\partial x}\right)+B\left(v\frac{\partial u}{\partial y}-u\frac{\partial v}{\partial y}\right)+\left(D-\frac{\partial A}{\partial x}-\frac{\partial B}{\partial y}\right)uv.$$

[29] *Jour. für Math.*, 104, 1889, 241-301.
[30] Vol. 2, libro IV, Cap. 4, 2.ª ed., 1915.

$M(v)$ es la expresión adjunta de $L(u)$ y $M(v) = 0$ es la ecuación diferencial adjunta. Inversamente, $L(u)$ es la adjunta de $M(v)$.

La importancia del teorema de Green reside en que puede ser usado para obtener soluciones de algunas ecuaciones diferenciales parciales. Así, ya que la ecuación elíptica puede ser siempre puesta en la forma

$$L(u) = \Delta u + a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y} + cu = 0,$$

entonces

$$M(v) = \Delta v - \frac{\partial(av)}{\partial x} - \frac{\partial(bv)}{\partial y} + cv = 0.$$

Sea v una solución de una ecuación adjunta que se hace infinita (logarítmicamente) en un punto arbitrario (ξ, η); esto es, se comporta como

$$v = U \log r + V,$$

donde r es la distancia de (ξ, η) a (x,y); U y V son continuas en el dominio R considerado; y U está normalizada de tal forma que $U(\xi, \eta) = 1$. Ahora exclúyase (ξ, η) del dominio de integración encerrándolo en un círculo. Entonces el teorema generalizado de Green proporciona, cuando el círculo se contrae a (ξ, η):

$$\begin{aligned} 2\pi u(\xi, \eta) = &-\iint vL(u)\,dx\,dy \\ &+ \int\left[v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n} + (a\cos(n,x) + b\cos(n,y))\cdot uv\right]ds \end{aligned} \tag{49}$$

donde n es positiva si está dirigida hacia el exterior del dominio; la integral simple se toma en sentido contrario a las manecillas del reloj sobre la frontera. Ya que u satisface $L(u) = 1$, si conocemos v y si u y $\partial u/\partial n$ están dadas (ambas no son arbitrarias) sobre la frontera, entonces hemos expresado u como una integral simple. La función v es llamada función de Green, aunque con frecuencia es añadida la

condición de que v se anula sobre la frontera de R en la definición de la función de Green. Han sido desarrolladas varias especializaciones y generalizaciones de este uso del teorema de Green.

7. Sistemas de ecuaciones en derivadas parciales

En el siglo XVIII las ecuaciones diferenciales del movimiento de fluidos fueron el primer sistema importante de ecuaciones diferenciales parciales. En el siglo XIX fueron creados tres sistemas fundamentales más —las ecuaciones de la dinámica de fluidos para medios viscosos, las ecuaciones del medio elástico y las ecuaciones de la teoría electromagnética.

La adquisición de las ecuaciones del movimiento de fluidos cuando la viscosidad está presente (como siempre está) siguió un camino tortuoso. Euler había proporcionado las ecuaciones del movimiento de un fluido que no es viscoso. Desde los tiempos de Lagrange ya había sido reconocida la diferencia esencial entre el movimiento de un fluido cuando existe un potencial de velocidades y cuando no. Llevado por una analogía formal con la teoría de la elasticidad y por la hipótesis de moléculas animadas por fuerzas repulsivas, Claude L. M. H. Navier (1785-1836), profesor de mecánica de la Ecole Polytechnique y de la Ecole des Ponts et Chaussées, obtuvo la ecuación básica en 1821 [31]. Las ecuaciones de Navier-Stokes, como son ahora identificadas, son

$$\begin{aligned}
\varrho \frac{Du}{Dt} &= \varrho X - \frac{\partial p}{\partial x} + \frac{1}{3}\mu \frac{\partial \theta}{\partial x} + \mu \Delta u \\
\varrho \frac{Dv}{Dt} &= \varrho Y - \frac{\partial p}{\partial y} + \frac{1}{3}\mu \frac{\partial \theta}{\partial y} + \mu \Delta v \\
\varrho \frac{Dw}{Dt} &= \varrho Z - \frac{\partial p}{\partial z} + \frac{1}{3}\mu \frac{\partial \theta}{\partial z} + \mu \Delta w \\
\theta &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z},
\end{aligned} \tag{50}$$

donde Δ tiene el significado usual; e es la densidad de un fluido; p

[31] *Mém. de l'Acad. des Sci., Paris* (2), 1827, 375-394.

la presión: *u*, *v* y *w* son las componentes de la velocidad de un fluido en cualquier *x*, *y*, *z* y *t*; *X*, *Y* y *Z* son las componentes de una fuerza externa; la constante μ, que depende de la naturaleza del fluido, es llamada coeficiente de viscosidad; y la derivada D/Dt tiene el significado explicado en el Capítulo 22, sección 8. Para un fluido incompresible, $\theta = 0$.

Las ecuaciones las obtuvo también Poisson en 1829 [32]. Después fueron encontradas de nuevo en 1845 sobre la base de la mecánica del continuo por George Gabriel Stokes (1819-1903), profesor de matemáticas de la universidad de Cambridge, en su ensayo «On the theories of the internal friction of fluids in motion» («Sobre las teorías de la fricción interna de fluidos de movimiento») [33]. Stokes se propuso explicar la acción de la fricción en todos los líquidos conocidos, lo que causa que el movimiento ceda al convertirse la energía cinética en calor. Los fluidos, en virtud de su viscosidad, se adhieren a las superficies de los sólidos y de esta manera ejercen fuerzas tangenciales sobre ellos.

La elasticidad fue fundada por Galileo, Hooke y Mariotte, y cultivada por los Bernoulli y Euler, aunque éstos trataron problemas específicos. Para resolverlos combinaron hipótesis *ad hoc* sobre cómo las vigas, varillas y láminas se comportaban bajo tensiones, presiones y pesos. La teoría propiamente dicha es una creación del siglo XIX. Desde el principio del siglo XIX un buen número de grandes hombres trabajaron persistentemente para obtener las ecuaciones que gobiernan el comportamiento de los medios elásticos, lo que incluye el aire. Estos científicos eran principalmente ingenieros y físicos; Cauchy y Poisson son los grandes matemáticos de entre ellos, aunque Cauchy era ingeniero por su formación.

Los problemas de elasticidad incluyen el comportamiento de cuerpos bajo tensión, donde se considera qué posición de equilibrio tomarán las vibraciones de los cuerpos cuando permanecen en movimiento por una perturbación inicial o mediante una fuerza aplicada continuamente, y, en el caso del aire y los cuerpos sólidos, la propagación de las ondas a través de ellos. El interés por la elasticidad en el siglo XIX fue intensificado por la aparición, alrededor de 1820, de una teoría ondulatoria de la luz, iniciada por el físico Thomas Young (1773-1829) y por Augustin-Jean Fresnel (1788-1827), un in-

---

[32] *Jour. de l'Ecole Poly.*, 13, 1831, 1-74.

[33] *Trans. Camb. Phil. Soc.*, 8, 1849, 287-319 = *Math. and Phys. Papers*, 1, 75-129.

geniero. La luz era concebida como un movimiento de ondas en el éter y el éter era pensado como un medio elástico. De aquí que la propagación de la luz a través del éter se convirtiera en un problema básico. Otro estímulo para el fuerte interés por la elasticidad en el siglo XIX lo fueron los experimentos (1787) de F. F. Chladni (1756-1827) sobre las vibraciones del vidrio y metal, en las que mostró las líneas nodales. Estas deben ser relacionadas con los sonidos proporcionados por, como ejemplo, una piel de tambor vibrando.

El trabajo para obtener las ecuaciones básicas de la elasticidad fue largo y lleno de escollos porque se sabía poco de la estructura interna o molecular de la materia; era difícil comprender cualquier problema físico. Las hipótesis hechas en cuanto a los cuerpos sólidos, el aire y el éter variaban de un autor a otro y estaban en disputa. En el caso del éter, que supuestamente penetraba los cuerpos sólidos, dado que la luz pasaba a través de algunos de ellos y era absorbida por otros, la relación de las moléculas del éter con las moléculas del cuerpo sólido también provocó grandes dificultades. No pretendemos seguir las teorías físicas de los cuerpos elásticos, nuestro entendimiento en este terreno no es claro ni aún hoy día.

Navier [34] fue el primero (1821) en investigar las ecuaciones generales del equilibrio y las vibraciones de los cuerpos elásticos. Se suponía que el material era isótropo y las ecuaciones contenían una constante única representando la naturaleza del sólido. En 1822, estimulado por el trabajo de Fresnel, Cauchy había creado otro enfoque de la teoría de la elasticidad [35]. Las ecuaciones de Cauchy contenían dos constantes para representar el material del cuerpo, y para un cuerpo isótropo son

$$
\begin{aligned}
(\lambda + \mu)\frac{\partial \theta}{\partial x} + \mu \Delta u &= \varrho \frac{\partial^2 u}{\partial t^2} \\
(\lambda + \mu)\frac{\partial \theta}{\partial y} + \mu \Delta v &= \varrho \frac{\partial^2 v}{\partial t^2} \\
(\lambda + \mu)\frac{\partial \theta}{\partial z} + \mu \Delta w &= \varrho \frac{\partial^2 w}{\partial t^2} \\
\theta = \frac{\partial u}{\partial x} &+ \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}.
\end{aligned}
\tag{51}
$$

[34] *Mém. de l'Acad. des Sci., Paris* (2), 7, 1827, 375-394.

[35] *Exercices de math.*, 1828 = *Œuvres* (2), 8, 195-226.

Aquí, u, v y w constituyen componentes de los desplazamientos, θ es llamado dilatación y λ y μ son las constantes del cuerpo o medio. Para un cuerpo general no isótropo las ecuaciones resultan bastante complicadas y sería inútil escribirlas con generalidad. Las ecuaciones están dadas por Cauchy [36].

El triunfo más espectacular del siglo XIX, con un enorme impacto sobre la ciencia y la tecnología, fue la derivación por Maxwell de las leyes del electromagnetismo [37]. Maxwell, utilizando las investigaciones eléctricas y magnéticas de numerosos predecesores, notablemente las de Faraday, introdujo la noción de corriente de desplazamiento —las ondas de radio son una forma de corriente de desplazamiento— y con esta noción formuló las leyes de la propagación de ondas electromagnéticas. Sus ecuaciones, más convenientemente enunciadas en la forma vectorial adoptada posteriormente por Oliver Heaviside, son cuatro en número y hacen intervenir el campo eléctrico de intensidad E, el campo magnético de intensidad H, la constante dieléctrica ε del medio, la permeabilidad magnética μ del medio y la densidad de carga p. Las ecuaciones son

$$\operatorname{rot} \mathbf{H} = \frac{1}{c}\frac{\partial(\varepsilon \mathrm{E})}{\partial t}, \qquad \operatorname{rot} \mathbf{E} = -\frac{1}{c}\frac{\partial(\mu \mathrm{H})}{\partial t} \tag{52}$$

$$\operatorname{div} \varepsilon\mathbf{E} = \varrho \quad \text{y} \quad \operatorname{div} \mu\mathbf{H} = 0. \tag{53}$$

[38]

Las dos primeras ecuaciones son las principales y suman seis ecuaciones diferenciales parciales escalares (no vectoriales). La corriente de desplazamiento es el término $\partial(\varepsilon E)/\partial t$.

Trabajando únicamente con estas ecuaciones, Maxwell predijo que las ondas electromagnéticas viajaban a través del espacio y a la velocidad de la luz. Sobre la base de la identidad de las dos velocidades, se atrevió a asegurar que la luz es un fenómeno electromagnético, una predicción que ha sido ampliamente confirmada desde entonces.

No se conocen métodos generales para resolver ninguno de los sistemas de ecuaciones anteriores. Sin embargo, los científicos del siglo XIX se dieron cuenta gradualmente de que en el caso de las

---

[36] *Exercices de math.*, 1828 = *Œuvres* (2), 253-277.

[37] *Phil. Trans.*, 155, 1865, 459-512 = *Scientific Papers*, 1, 526-597.

[38] Para el significado de rot y div véase el Cap. 32, sec. 5.

ecuaciones diferenciales parciales, ya fueran simples ecuaciones o sistemas, las soluciones generales no son tan claramente útiles como las soluciones para problemas específicos donde las condiciones iniciales y de frontera están dadas, y el trabajo experimental también puede ayudar haciendo suposiciones simplificadas útiles. Los escritos de Fourier, Cauchy y Riemann llevaron más lejos esta tendencia. El trabajo sobre la solución de los muchos problemas de valor inicial y de frontera a los que especializaciones de estos sistemas dieron origen es enorme y la gran mayoría de los matemáticos del siglo se dedicaron a dichos problemas.

8. Teoremas de existencia

Así como los matemáticos de los siglos XVIII y XIX crearon un vasto número de tipos de ecuaciones diferenciales, se encontraron con que los métodos para resolver muchas de estas ecuaciones no estaban prontos. De la misma manera que en el caso de las ecuaciones polinominales, donde los esfuerzos por resolver ecuaciones de grado mayor que cuatro fallaron, y Gauss se volvió hacia la prueba de la existencia de una raíz (cap. 25, sec. 2), así en el caso de las ecuaciones diferenciales el fracaso en encontrar soluciones explícitas, las que por supuesto demuestran *ipso facto* la existencia, forzaron a los matemáticos a volverse hacia demostraciones de la existencia de las soluciones. Tales pruebas, a pesar de que no exhiben una solución, ni la dan de una manera útil, cumplen varios propósitos. Las ecuaciones diferenciales fueron en casi todos los casos la formulación matemática de problemas físicos. No se garantizaba que las ecuaciones matemáticas pudieran ser resueltas; una demostración de la existencia de una solución podría asegurar, al menos, que la búsqueda de una solución no representaba un imposible. La prueba de existencia también contestaría la pregunta: ¿qué debemos saber acerca de una situación física dada, esto es, qué condiciones iniciales y de frontera aseguran una solución, preferentemente única? Otros objetivos, tal vez no previstos al inicio del trabajo sobre teoremas de existencia, fueron reconocidos de inmediato. ¿La solución cambia continuamente con las condiciones iniciales, o algún fenómeno enteramente nuevo aparece cuando las condiciones de frontera o iniciales son variadas ligeramente? Así, una órbita parabólica que se obtiene para un valor de la velocidad inicial de un planeta puede

cambiar a una órbita elíptica como consecuencia de un pequeño cambio de la velocidad inicial. Tal diferencia en la órbita es físicamente de lo más significativo. Más aún, algunas de los métodos de solución, tal como el uso del principio de Dirichlet, o el teorema de Green, presuponían la existencia de una solución particular. La existencia de estas soluciones particulares no había sido establecida.

Antes de dar algunas breves indicaciones del trabajo sobre teoremas de existencia, puede ser útil señalar una clasificación de las ecuaciones diferenciales parciales que fue de hecho bosquejada bastante tarde en este siglo. A pesar de que Laplace y Poisson habían realizado algunos esfuerzos hacia la clasificación, reduciendo estas ecuaciones a formas normales, la clasificación introducida por Du Bois Reymond se ha hecho obligada. En 1889 [39], Du Bois Reymond clasificó, por medio de sus características, las ecuaciones lineales homogéneas más generales de segundo orden

$$R\frac{\partial^2 u}{\partial x^2} + S\frac{\partial^2 u}{\partial x\,\partial y} + T\frac{\partial^2 u}{\partial y^2} + P\frac{\partial u}{\partial x} + Q\frac{\partial u}{\partial y} + Zu = 0, \qquad (54)$$

donde los coeficientes son funciones de x e y y éstas, y sus primeras y segundas derivadas, son continuas, por medio de las características (cap. 22, sec. 7). Las proyecciones de las curvas características dentro del plano xy (estas proyecciones también son llamadas características) satisfacen

$$T\,dx^2 - S\,dx\,dy + R\,dy^2 = 0.$$

Las características son imaginarias, reales y distintas, o reales y coincidentes de acuerdo con que sea

$$TR - S^2 > 0, \qquad TR - S^2 < 0, \qquad TR - S^2 = 0.$$

Du Bois Reymond llamó estos casos elíptico, hiperbólico y parabólico, respectivamente, y señaló que introduciendo nuevas variables reales independientes

$$\xi = \phi(x, y), \qquad \eta = \psi(x, y),$$

[39] *Jour. für Math.*, 104, 1889, 241-301.

la ecuación anterior siempre puede ser transformada en uno de los tres tipos de formas normales

$$\text{(a)}\quad R'\left(\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \eta^2}\right)+P'\,\frac{\partial u}{\partial \xi}+Q'\,\frac{\partial u}{\partial \eta}+Zu=0$$

$$\text{(b)}\quad S'\,\frac{\partial^2 u}{\partial \xi\,\partial \eta}+P'\,\frac{\partial u}{\partial \xi}+Q'\,\frac{\partial u}{\partial \eta}+Zu=0$$

$$\text{(c)}\quad R'\,\frac{\partial^2 u}{\partial \xi^2}+P'\,\frac{\partial u}{\partial \xi}+Q'\,\frac{\partial u}{\partial \eta}+Zu=0,$$

respectivamente. Las dos familias $\phi(x,y)=$ const. y $\psi(x,y)=$ const. son las ecuaciones de dos familias de características.

Las condiciones suplementarias que pueden ser impuestas difieren para los tres tipos de ecuaciones. En el caso elíptico (a) se considera un dominio acotado del plano xy y se especifica el valor de u en la frontera (o una condición equivalente) y se pregunta por el valor de u en el dominio. Para el problema de valor inicial de la ecuación diferencial hiperbólica (b) se debe especificar u y $\partial u/\partial n$ sobre alguna curva inicial. También puede haber condiciones en la frontera. Las condiciones iniciales apropiadas para el caso parabólico (c) no fueron especificadas en aquel momento, aunque ahora es sabido que una condición inicial y condiciones en la frontera pueden ser impuestas. Esta clasificación de las ecuaciones diferenciales parciales se extendió a ecuaciones con más variables independientes, ecuaciones de orden superior, y a sistemas. A pesar de que la clasificación y las condiciones suplementarias no eran conocidas a principios de siglo, gradualmente los matemáticos advirtieron las distinciones y éstas figuraron en los teoremas de existencia que fueron capaces de demostrar.

El trabajo sobre los teoremas de existencia se convirtió en una actividad de mayor relevancia con Cauchy, quien insistió en que la existencia puede ser establecida frecuentemente cuando aún no esté disponible una solución específica. En una serie de artículos [40], Cauchy hizo notar que cualquier ecuación diferencial parcial de orden superior a uno puede ser reducida a un sistema de ecuaciones diferenciales parciales, y se ocupó de la existencia de una solución

[40] *Comp. Rend.*, 14, 1842, 1020-1025 = *Œuvres* (1), 6 461-467 y *Comp. Rend.*, 15, 1842, 44-59, 85-101, 131-138 = *Œuvres* (1), 7, 17-33, 33-49, 52-58.

para el sistema. Llamó a su método el *calcul des limites* (cálculo de límites), pero es conocido ahora como el método de las funciones mayorantes. La esencia del método es mostrar que una serie de potencias en las variables independientes con un dominio definido de convergencia satisface el sistema de ecuaciones. Ilustraremos el método en relación con el trabajo de Cauchy sobre ecuaciones diferenciales ordinarias (cap. 29, sec. 4). Su teorema cubre únicamente el caso de coeficientes analíticos en las ecuaciones y condiciones iniciales analíticas.

Para obtener una idea concreta del trabajo de Cauchy, vamos a considerar lo que implica para la ecuación de segundo orden en dos variables independientes

$$r = f(z, x, y, p, q, s, t) \tag{55}$$

donde, como es usual, $r = \partial^2 z/\partial x^2$ y f es analítica en todas las variables. En este caso se debe especificar sobre la recta inicial $x = 0$ que sea

$$z(0, y) = z_0(y), \qquad \frac{\partial z}{\partial x}(0, y) = z_1(y),$$

donde z_0 y z_1 son analíticas. (La recta inicial puede ser reemplazada por una curva, en cuyo caso $\partial z/\partial x$ debe ser sustituida por $\partial z/\partial n$). Si las condiciones anteriores se cumplen, entonces la solución $z = z(x,y)$ existe y es única, y analítica en algún dominio empezando en la línea inicial.

El trabajo de Cauchy sobre sistemas lo realizó independientemente, y de alguna manera mejorado, Sophie Kowalewsky (1850-1891) [41], que fue alumna de Weierstrass y prosiguió sus ideas. La Kowalewsky es una de las pocas mujeres matemáticas distinguidas. En 1816, Sophie Germain (1776-1831) obtuvo un premio por un ensayo sobre elasticidad que le otorgó la Academia Francesa. Kowalewsky, ganó asimismo el premio de la Academia de París, con un trabajo de 1888 sobre la integración de las ecuaciones del movimiento para un cuerpo sólido rotando alrededor de un punto fijo; en 1889 llegó a ser profesor de matemáticas en Estocolmo. Las demostraciones de Cauchy y Kowalewsky fueron mejoradas posteriormente por Goursat [42].

---

[41] *Jour. für Math.*, 80, 1875, 1-32.
[42] *Bull. Soc. Math. de France*, 26, 1898, 129-134.

Si, en lugar de (55), la ecuación de segundo orden dada es de la forma

$$G(z, x, y, p, q, r, s, t) = 0, \tag{56}$$

es necesario resolverla en r antes de que pueda ser puesta en la forma (55). Para considerar un caso simple, pero esencial, si la ecuación es

$$G = A\frac{\partial^2 z}{\partial x^2} + 2B\frac{\partial^2 z}{\partial x\,\partial y} + C\frac{\partial^2 z}{\partial y^2} + D\frac{\partial z}{\partial x} + E\frac{\partial z}{\partial y} + Fz = 0,$$

donde A, B,..., F son funciones de x e y, entonces $\partial G/\partial r$ no debe ser 0 para despejar r. En el caso $\partial G/\partial r = 0$, la solución del problema de Cauchy puede no existir, y cuando lo hace no es única. En el caso de tres o más variables independientes (consideremos tres), y si la ecuación es escrita como

$$\sum_{i,k} A_{ik}\frac{\partial^2 u}{\partial x_i\,\partial x_k} + \sum_i B_i\frac{\partial u}{\partial x_i} + Cu = f, \tag{57}$$

donde los coeficientes son funciones de las variables independientes x_1, x_2 y x_3, entonces el caso excepcional sucede cuando la superficie inicial S satisface la ecuación diferencial parcial de primer orden

$$\sum A_{ik}\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_k} = 0. \tag{58}$$

A lo largo de tales superficies las dos soluciones de (57) pueden ser tangentes e incluso tener un contacto de orden superior. Esta propiedad es la misma para las curvas características de la ecuación de primer orden $f(x, y, u, p, q) = 0$ (cap. 22, sec. 5), y así estas superficies S también son llamadas características. Físicamente, las superficies S son frentes de onda.

Esta teoría de las características para el caso de dos variables independientes era conocida por Monge y por André-Marie Ampère (1775-1836). Su extensión para ecuaciones de segundo orden en más de dos variables independientes lo efectuó por primera vez Albert Victor Bäcklund (1845-1922) [43], pero este caso no fue ampliamente conocido hasta que no lo rehízo Jules Beudon [44].

[43] *Math. Ann.* 13, 1878, 411-428.
[44] *Bull. Soc. Math. de France*, 25, 1897, 108-120.

En su *Leçons sur la propagation des ondes* (*Lecciones sobre la propagación de ondas,* 1903), Jacques Hadamard (1865-1963), el gran matemático francés de este siglo, generalizó la teoría de las características para ecuaciones diferenciales parciales de segundo orden en las variables dependientes ξ, η y ζ y las variables independientes x_1, $x_2,\ldots, x_n$. El problema de Cauchy para este sistema es: dados los valores de ξ, η, ζ y $\partial\xi/\partial X_n$, $\partial\eta/\partial X_n$ y $\partial\zeta/\partial x_n$ sobre la «superficie» M_{n-1} de dimensión $n-1$, encontrar las funciones ξ, η y ζ. Los valores de ξ, η, ζ y $\partial\xi/\partial x_n$, $\partial\eta/\partial x_n$ y $\partial\zeta/\partial x_n$ sobre la «superficie» lados a menos que M_{n-1} satisfaga una ecuación diferencial parcial de primer orden de sexto grado, digamos $H = 0$. Todas las «superficies» que cumplen $H = 0$ son «superficies» características. De acuerdo con la teoría de ecuaciones diferenciales parciales de primer orden, la ecuación diferencial $H = 0$ tiene líneas (curvas) características definidas por

$$\frac{dx_1}{\partial H/\partial P_1} = \frac{dx_2}{\partial H/\partial P_2} = \ldots = \frac{dx_{n-1}}{\partial H/\partial P_{n-1}},$$

donde P_1, $P_2,\ldots, P_{n-1}$ son las derivadas parciales de x_n con respecto a x_1, $x_2,\ldots, x_{n-1}$ calculadas a lo largo de la «superficie» M_{n-1}. Estas líneas son llamadas las bicaracterísticas del sistema original de segundo grado. En la teoría de la luz son los rayos.

Las características juegan ahora un papel vital en la teoría de ecuaciones en derivadas parciales. Por ejemplo, Darboux [45] proporcionó un método poderoso para integrar ecuaciones diferenciales parciales de segundo orden en dos variables independientes que se basa en la teoría de las características: transforma el problema en la integración de una o más ecuaciones diferenciales ordinarias e incluye los métodos de Monge, Laplace y otros.

Otro caso de teoremas de existencia trataba el problema de Dirichlet, esto es, estableciendo la existencia de una solución para $\Delta V = 0$, ya fuera directamente o por medio del principio de Dirichlet. La primera prueba de existencia del problema de Dirichlet en dos dimensiones (pero no del principio de Dirichlet de minimizar la integral de Dirichlet) fue proporcionado por Hermann Amandus Schwarz (1843-1921), alumno de Weierstrass, a quien sucedió en Berlín en 1892, quien le sugirió el problema. Bajo hipótesis generales

---

[45] *Ann. de l'Ecole Norm. Sup.* (1), 7, 1870, 175-180.

acerca de la curva frontera y usando un proceso llamado el procedimiento de la alternativa [46], demostró la existencia de una solución [47].

En el mismo año, 1870, Carl G. Neumann proporcionó otra demostración de la existencia de una solución al problema de Dirichlet en tres dimensiones [48], usando el método de medias aritméticas, a pesar de que él mismo tampoco usó el principio de Dirichlet [49]. La principal exposición de sus ideas está en su *Vorlesungen uber Riemann's Theorie des Abel'schen Integrale (Lecciones sobre la teoría de Riemann de las integrales abelianas)* [50].

Entonces Poincaré [51] utilizó el *méthode de balayage* (el método de «barrido»), que ataca el problema construyendo una sucesión de funciones no armónicas en el dominio *R* pero tomando los valores de contorno correctos, haciendo las funciones más y más armónicas.

Finalmente, David Hilbert reconstruyó el método del cálculo de variaciones de Thomson y Dirichlet, y estableció el *principio* de Dirichlet como un método para probar la existencia de una solución del problema de Dirichlet. En 1899 [52], Hilbert demostró que bajo las condiciones adecuadas sobre una región, los valores en la frontera y las funciones admisibles *U*, se cumple el principio de Dirichlet, e hizo del principio de Dirichlet una herramienta poderosa en la teoría de funciones. En otras publicación respecto a trabajos realizados en 1901 [53], Hilbert proporcionó condiciones más generales.

La historia del principio de Dirichlet es notable. Green, Dirichlet, Thomson y otros de sus contemporáneos lo consideraban como un método bien fundado, utilizándolo libremente. Más adelante, Riemann, en su teoría de funciones complejas, demostró que era un instrumento extraordinario para llegar a resultados importantes. To-

---

[46] El método de Schwarz está esbozado en el libro de Felix Klein: *Vorlesungen uber die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea (reimp.), 1950, 1, p. 265 y está dado en su totalidad en: A. R. Forsyth, *Theory of Functions*, Dover (reimp.), 1965, 2, Cap. 17. Esta última fuente incluye muchas referencias.

[47] *Monatsber. Berliner Akad.*, 1870, 767-795 = *Ges. Math. Abh.*, 2, 144-171.

[48] *Königlich Sachsischen Ges. der Wiss. zu Leipzig*, 1870, 49-56, 264-321.

[49] El método está descrito en: O. D. Kellogg, *Foundations of Potential Theory*, Julius Springer, 1929, 281 sigs.

[50] 2.ª ed., 1884, 238 sigs.

[51] *Amer. Jour. of Math.*, 12, 1890, 211-294 = *Œuvres*, 9, 28-113.

[52] *Jahres. der Deut. Math.-Verein.*, 8, 1900, 184-188 = *Ges. Abh.*, 3, 10-14.

[53] *Math. Ann.*, 59, 1904, 161-186 = *Ges. Abh.*, 3, 15-37.

dos estos hombres eran conscientes de que la cuestión fundamental de la existencia no estaba aclarada, aún antes de que Weierstrass anunciara su crítica de 1870, que desacreditó el método durante varias décadas. Entonces Hilbert rescata el principio que fue usado y difundido en este siglo. Si el progreso que se había logrado con el uso de este principio hubiera esperado el trabajo de Hilbert, gran parte del trabajo sobre teoría del potencial y teoría de funciones del siglo XIX se hubiera perdido.

La ecuación de Laplace $\Delta V = 0$ es la forma básica de las ecuaciones diferenciales elípticas. Se establecieron muchos más teoremas de existencia para más ecuaciones diferenciales elípticas generales, tales como

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + a\,\frac{\partial u}{\partial x} + b\,\frac{\partial u}{\partial y} + cu = 0. \qquad (59)$$

La variedad de tales teoremas es grande. Mencionaremos sólo un resultado clave. La existencia y unicidad de una solución de esta ecuación (el valor de la solución está prescrito en la frontera) fue demostrado por Picard [54] para dominios de área suficientemente pequeña. El resultado ha sido extendido a más variables, dominios más amplios y en otros varios aspectos por Picard y otros. Picard también estableció [55] que toda ecuación de la forma anterior (y aun un poco más general) cuyos coeficientes son funciones analíticas, posee sólo soluciones analíticas dentro del dominio en el que la solución es buscada, aunque la solución tome valores no analíticos en la frontera del dominio.

Los teoremas discutidos hasta aquí han tratado generalmente de ecuaciones diferenciales y datos iniciales o de frontera analíticos. Sin embargo, tales condiciones son demasiado restrictivas para las aplicaciones, porque los datos físicos dados pueden no ser analíticos. Otra clase importante de teoremas trata de condiciones menos restringidas. Daremos un solo ejemplo. El método de Riemann, que se aplica a la ecuación hiperbólica, se apoya en la existencia de su función característica v y, como señalamos, la existencia de v no fue establecida por Riemann.

---

[54] *Comp. Rend.*, 107; 1888, 939-941, *Jour. de Math.* (4), 6, 1890, 145-210, *Jour. de Math.* (5), 2, 1896, 295-304.

[55] *Jour. de l'Ecole Poly.*, 60, 1890, 89-105.

Para este caso hiperbólico (véase [40]), Du Bois Reymond buscó en 1889 las condiciones adecuadas y obtuvo resultados [56] que, expresados para el caso donde x = const. e y = const. son las características, dicen: dadas dos funciones continuas u y $\partial u/\partial n$ a lo largo de una curva AB que no es cortada más de una vez por cualquier línea característica, existe entonces una y solamente una solución u de la ecuación diferencial que toma los valores dados de u y $\partial u/\partial n$ a lo largo de AB. Esta solución está definida en el rectángulo determinado por las características que pasan por A y B. Si en lugar de esto se dan los valores de una función continua u sobre dos segmentos de características que colindan uno con otro, entonces de nuevo u está determinada únicamente en el rectángulo determinado por las características. En términos de x, y y u como coordenadas espaciales, el primer resultado afirma que la superficie $u(x,y)$ pasa a través de una curva dada en el espacio y con una inclinación dada. El segundo resultado significa que la solución o superficie está encerrada en el espacio determinado por dos curvas que se cortan. Para condiciones iniciales continuas u y $\partial u/\partial n$ (y u en el segundo caso), las soluciones serán regulares o bien satisfacen la ecuación diferencial en todo punto de los rectángulos descritos anteriormente. Las discontinuidades en u y $\partial u/\partial n$ se propagarán a lo largo de las características en el rectángulo.

Una gran cantidad de trabajo sobre la existencia de los valores característicos para $\Delta u + k^2 u = 0$, considerada en el dominio D, fue realizada en la segunda mitad del siglo. El resultado principal es que, para un dominio dado y bajo cualquiera de las condiciones en la frontera $u = 0$, $\partial u/\partial n = 0$, $\partial u/\partial n + hu = 0$ ($h > 0$ cuando la normal está dirigida fuera del dominio), existe siempre un número infinito de valores discretos de k^2, y que para cada uno de ellos existe una solución. En dos dimensiones, las vibraciones de una membrana fijada a lo largo de su frontera ilustran este teorema. Los valores de k son las frecuencias de la infinidad de vibraciones puramente armónicas. Las soluciones correspondientes proporcionan la deformación de la membrana al llevar a cabo sus oscilaciones características.

El primer gran paso fue la demostración de Schwarz [57] de la existencia de la primera función característica de

$$\Delta u + \xi f(x, y) u = 0,$$

[56] *Jour. für Math.*, 104, 1889, 241-301.

[57] *Acta Soc. Fennicae*, 15, 1885, 315-362 = *Ges. Math. Abh.*, 1, 223-269.

esto es, la existencia de una U_1 tal que

$$\Delta U_1 + k_1^2 f(x, y) U_1 = 0$$

y $U_1 = 0$ sobre la frontera del dominio considerado. Su método proporcionó un procedimiento para encontrar la solución y permitió el cálculo de k_1^2. Picard [58] estableció entonces la existencia del segundo valor característico k_2^2.

En su artículo de 1885, Schwarz también demostró que cuando el dominio varía continuamente, el valor de k_1^2, el primer número característico, también varía continuamente; y al mismo tiempo que el dominio se hace más pequeño, k_1^2 se incrementa sin límite. Así, una membrana pequeña da origen a un primer armónico superior.

En 1894, Poincaré [59] demostró la existencia y las propiedades esenciales de los valores característicos de

$$\Delta u + \lambda u = f, \tag{60}$$

con λ complejo, en un dominio tridimensional acotado, con $u = 0$ sobre la frontera. La existencia de u fue demostrada mediante una generalización del método de Schwarz. Acto seguido demostró que $u(\lambda)$ es una función meromorfa de la variable compleja λ y que los polos son reales; éstos son justamente los valores propios λ_n. Entonces obtuvo las soluciones características U_i; esto es,

$$\begin{aligned} \Delta U_i + k_i^2 U_i &= 0 \quad \text{(en el interior)} \\ U_i &= 0 \quad \text{(en el borde).} \end{aligned}$$

Los k_1^2 son los números característicos (valores propios) y determinan las frecuencias de las soluciones características respectivas.

Físicamente, el resultado de Poincaré tiene el significado siguiente. La función f en (60) puede pensarse como una fuerza aplicada. Las oscilaciones libres de un sistema mecánico son aquellas en que las oscilaciones forzadas degeneran y se hacen infinitas. De hecho, (60) es la ecuación de un sistema oscilante excitado por una fuerza variando periódicamente de amplitud f; y las soluciones caracterís-

[58] *Comp. Rend.*, 117, 1893, 502-507.

[59] *Rendiconti del Circolo Matematico di Palermo*, 8, 1894, 57-155 = *Œuvres*, 9, 123-196.

ticas son las oscilaciones libres del sistema, que, una vez excitadas, continúan indefinidamente. Las frecuencias de las oscilaciones libres, que son proporcionales a k_i, se calculan mediante el método de Poincaré como los valores $\sqrt{\lambda_i}$ para los cuales la oscilación forzada u se hace infinita.

Al final del siglo, la teoría sistemática de problemas de frontera y de valores iniciales para ecuaciones diferenciales parciales, que datan del ensayo fundamental de Schwarz de 1885, aún era muy joven. El trabajo en esta área se extendió rápidamente en el siglo XX.

Bibliografía

Bacharach, Max: *Abriss der Geschichte der Potentialtheorie*, Vandenhoeck y Ruprecht, 1883.

Burkhardt, H.: «Entwicklungen nach oscillirenden Functionen und Integration der Differentialgleichungen der mathematischen Physik», *Jahres. der Deut. Math.-Verein.*, vol. 10, 1908, 1-1804.

Burkhardt, H. y Franz Meyer, W.: «Potentialtheorie», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, II A7b, 464-503.

Cauchy, Augustin Louis: *Œuvres complètes*, Gauthier-Villars, 1890 (2), 8.

Fourier, Joseph: *The Analytical Theory of Heat* (1822). Dover, reimpresión de la traducción inglesa, 1955.

— : *Œuvres*, 2 vols., Gauthier-Villars, 1888-1890; Georg Olms (reimpresión), 1970.

Green, George: *Essay on the Application on Mathematical Analysis to the Theories of Electricity and Magnetism*, 1828, reimpresa por Wezata Melins Aktiebolang, 1958; también en Ostwald's Klassiker 61 (en alemán), Wilhelm Engelmenn, 1895.

— : *Mathematical Papers*, Macmillan, 1871; Chelsea (reimpresión), 1970.

Heine, Eduard: *Handbuch der Kugelfunktionen*, 2 vols., 1878-1881, Physika Verlag (reimpresión), 1961.

Helmholtz, Hermann von: «Theorie der Luftschwingungen in Röhren mit offenen Enden», *Ostwald's Klassiker der exakten Wissenschaften*, Wilhelm Engelmenn, 1896.

Klein, Felix: *Vorlesungen uber die Entwicklung der Mathematik im 19. Jahrhundert*, 2 vols., Chelsea (reimpresión), 1950.

Langer, R. E.: «Fourier Series, the Genesis and Evolution of a Theory», *Amer. Math. Monthly* 54, Part II, 1947, 1-86.

Pockels, Friederich: *Über die partielle Differentialgleichung* $\Delta u + k^2\mu = 0$, B. G. Teuber, 1891.

Poincaré, Henri: *Œuvres*, Gauthier-Villars, 1954, 9.

Rayleigh, lord (Strutt, John William): *The Theory of Sound*, 2.ª ed., 2 vols., Dover (reimpresión), 1945.
Riemann, Bernhard: *Gesammelte mathematische Werke*, 2.ª ed., Dover (reimpresión), 1953, pp. 156-211.
Sommerfield, Arnold: «Randwertaufgaben in de Theorie der partiellen Differentiallgleichungen», *Encyk. der math. Wiss.*, B. G. Teubner, 1899-1916, II A7c, 504-570.
Todhunter, Isaac y Pearson, Karl: *A History of the Theory of Elascticity*, 2 vols., Dover (reimpresión), 1960.
Whittaker, sir Edmund: *History of the Theories of Aether and Electricity*, ed. rev., Thomas Nelson and Sons, 1951, vol. I.

Capítulo 29
LAS ECUACIONES DIFERENCIALES ORDINARIAS EN EL SIGLO XIX

> La física no solamente nos ha dado la ocasión de resolver problemas (...) sino que también nos ha hecho presentir la solución.
>
> HENRI POINCARÉ

1. Introducción

Las ecuaciones diferenciales ordinarias surgieron en el siglo XVIII como una respuesta directa a problemas físicos. Al estudiar fenómenos físicos complicados, notablemente en los trabajos sobre la cuerda vibrante, los matemáticos llegaron a las ecuaciones en derivadas parciales. En el siglo XIX, los papeles que jugaron estas dos materias se invirtieron. El esfuerzo para resolver ecuaciones diferenciales parciales por el método de separación de variables condujo al problema de resolver ecuaciones diferenciales ordinarias. Más aún, puesto que las ecuaciones diferenciales parciales fueron expresadas en diversos sistemas de coordenadas, las ecuaciones diferenciales ordinarias resultantes que se presentaron eran extrañas e insolubles en forma explícita. Los matemáticos recurrieron a soluciones en series infinitas, que son actualmente conocidas como funciones especiales, de funciones trascendentes superiores como opuestas a las funciones trascendentes elementales tales como sen x, e^x y log x.

Después de mucho trabajo sobre la amplia variedad de ecuaciones diferenciales ordinarias, fueron dedicados algunos estudios teóricos profundos a ciertos tipos de tales ecuaciones. Estas investiga-

ciones teóricas también diferenciaron el trabajo efectuado en el siglo XIX del que se hizo en el XVIII. Las contribuciones del nuevo siglo fueron tan vastas que, como en el caso de las ecuaciones diferenciales parciales, no podemos esperar revisar todos los desarrollos principales. Nuestros temas sólo representan un ejemplo de lo que fue creado durante el siglo.

2. Soluciones con series y funciones especiales

Como acabamos de observar, para resolver las ecuaciones diferenciales ordinarias que resultaron del método de separación de variables aplicadas a ecuaciones diferenciales parciales, los matemáticos, sin preocuparse gran cosa por la existencia y forma que las soluciones deberían tener, retomaron el método de las series infinitas (cap. 21, sec. 6). De las ecuaciones diferenciales ordinarias que resultaron de la separación de variables, la más importante es la ecuación de Bessel.

$$x^2y'' + xy' + (x^2 - n^2)y = 0, \tag{1}$$

donde n, un parámetro, puede ser complejo y x también puede ser complejo. Sin embargo, para Friedrich Wilhelm Bessel (1784-1846), matemático y director del observatorio astronómico de Königsberg, n y x eran reales. Casos especiales de esta ecuación datan de 1703, cuando Jacques Bernoulli los mencionó como una solución particular en una carta a Leibniz, y más tarde en trabajos más extensos de Daniel Bernoulli y Euler (cap. 21, secs. 4 y 6 y cap. 22, sec. 3). Asimismo, aparecieron casos especiales en los escritos de Fourier y Poisson. El primer estudio sistemático de las soluciones de esta ecuación fue hecho por Bessel [1] mientras trabajaba sobre el movimiento de los planetas. La ecuación tiene dos soluciones independientes para cada n, denotadas hoy en día por $J_n(x)$ y $Y_n(x)$, llamadas las funciones de Bessel de la primera y segunda clase, respectivamente. Bessel, cuyo trabajo sobre la ecuación se remonta a 1816, proporcionó primero la relación integral (para n entero)

$$J_n(x) = \frac{1}{2\pi}\int_0^{2\pi} \cos(nu - x \operatorname{sen} u)\, du$$

[1] *Abh. König. Akad. der Wiss. Berlin*, 1824, 1-52, pub. 1826 = *Werke*, 1, 84-109.

(escribió I_k^h y su k es nuestra x). Bessel también obtuvo la serie

$$J_n(x)=\frac{x^n}{2^n\Gamma(n+1)}\left\{1-\frac{x^2}{2^2\cdot 1\,!\,(n+1)}+\frac{x^4}{2^4\cdot 2!(n+1)(n+2)}-\dots\right\}. \qquad (2)$$

En 1818, Bessel demostró que $J_0(x)$ tiene un número infinito de ceros reales. En el artículo de 1824, Bessel también proporcionó la fórmula recursiva (para n entero)

$$xJ_{n+1}(x) - 2nJ_n(x) + xJ_{n-1}(x) = 0,$$

y muchas otras relaciones concernientes a la función de Bessel de primera clase. La generalización de la función de Bessel $J_n(x)$ a n y x complejas fue hecha por varios autores [2], con (2) como forma correcta.

Debido a que han de existir dos soluciones independientes para una ecuación de segundo orden, muchos matemáticos las buscaron. Cuando n no es un entero, esta segunda solución es $J_{-n}(x)$. Para n entero, la segunda solución fue proporcionada por Carl G. Neumann [3]. Sin embargo, la forma más comúnmente adoptada hoy día fue dada por Hermann Hankel (1839-1873) [4], a saber,

$$Y_n(z)=\sum_{r=0}^{\infty}\frac{(-1)^r[(1/2)z]^{n+2r}}{r!\,(n+r)!}\left\{2\log\left(\frac{z}{2}\right)+2\gamma-\sum_{m=1}^{n+r}\frac{1}{m}-\sum_{m=1}^{r}\frac{1}{m}\right\}$$

$$-\sum_{r=0}^{n-1}\frac{[(1/2z]^{-n+2r}(n-r-1)!}{r!},$$

donde γ es la constante de Euler. Neumann [5] también proporcionó la expansión de una función analítica $f(z)$, a saber,

$$f(z) = \alpha_0 J_0(z) + \alpha_1 J_1(z) + \alpha_2 J_2(z) + \dots,$$

donde las α_i son constantes y pueden ser determinadas.

---

[2] Principalmente por Eugen C. J. Lommel (1837-1899) en su *Studien über die Bessel'schen Funktionen* (1868).

[3] *Theorie der Bessel'schen Funktionen*, 1867, 41.

[4] *Math. Ann.*, 1, 1869, 467-501.

[5] *Jour. für Math.*, 67, 1867, 310-314.

Muchos matemáticos, trabajando por lo general en mecánica celeste, llegaron independientemente a las funciones de Bessel y a cientos de otras relaciones y expresiones para estas funciones. Alguna idea de la vasta literatura sobre dichas funciones puede obtenerse revisando *A treatise on the Theory of Bessel Functions (Un tratado sobre la teoría de las funciones de Bessel)* [6] de G. N. Watson.

Los polinomios de Legendre o funciones esféricas de una variable y las funciones de esféricas de superficie, que son funciones de dos variables, habían sido ya presentadas por Legendre y Laplace (cap. 22, sec. 4). Los polinomios de Legendre satisfacen la ecuación diferencial de Legendre.

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0.$$

Esta ecuación, como la conocemos, resulta de la separación de variables aplicada a la ecuación del potencial expresada en coordenadas esféricas. En 1833, Robert Murphy (murió en 1843), miembro de la universidad de Cambridge, escribió un texto, *Elementary Principles of the Theories of Electricity, Heat and Molecular Actions (Principios elementales de las teorías de electricidad, calor y acciones moleculares).* En él reunió algunos viejos resultados sobre los polinomios de Legendre y obtuvo algunos nuevos. Puesto que los principales resultados eran conocidos, no presentaremos los detalles del trabajo de Murphy, excepto para puntualizar que fue sistemático y demostró que «cualquier» función $f(x)$ puede ser desarrollada en términos de las $P_n(x)$ aplicando la integración término a término y la propiedad de la ortogonalidad (teorema integral).

Heine [7], tratando el problema del potencial para el exterior de un elipsoide de revolución y para la concha entre elipsoides de revolución cofocales (cap. 28, sec. 5), introdujo armónicos esféricos de segunda clase, denotados comúnmente por $Q_n(x)$, que dan una segunda solución independiente de la ecuación de Legendre. Las funciones de Legendre, como las funciones de Bessel, habían sido extendidas a n complejos y x complejos y habían sido obtenidas [8] gran número de representaciones alternativas y relaciones entre ellas.

---

[6] Cambridge University Press, 2.ª ed., 1944.

[7] *Jour. für Math.*, 26, 1843, 185-216.

[8] Véase, por ejemplo, E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, 1931, Chelsea (reim.), 1955.

El estudio de las funciones especiales que surgen como soluciones en serie de ecuaciones diferenciales ordinarias fue llevado más lejos por Gauss en su famoso ensayo de 1812 sobre las series hipergeométricas [9]. En este ensayo, Gauss no hizo uso de la ecuación diferencial

$$x(1-x)y'' + \{\gamma - (\alpha + \beta + 1)x\}y' - \alpha\beta y = 0,$$

pero sí lo hizo en material inédito [10]. Por supuesto, la ecuación y la solución en serie

$$F(\alpha, \beta, \gamma; x) = 1 + \frac{\alpha \cdot \beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2 \cdot \gamma(\gamma+1)} x^2 + \ldots \qquad (3)$$

eran ya conocidas, porque Euler las había estudiado (cap. 21, sec. 6). Gauss reconoció que para valores especiales de α, β y γ la serie incluía casi todas las funciones elementales conocidas hasta entonces y muchas funciones trascendentes superiores tales como las funciones de Bessel y las funciones esféricas. Además de demostrar gran número de propiedades de la serie, Gauss estableció la famosa relación

$$F(\alpha, \beta, \gamma; 1) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}.$$

También estableció la convergencia de la serie (*cf.* cap. 40, sec. 5). La notación $F(\alpha, \beta, \gamma; x)$ se debe a Gauss.

Lamé introdujo otra clase de funciones especiales [11]. En el cap. 28, sec. 5, señalamos que Lamé, trabajando sobre una distribución de temperatura en estado estacionario en un elipsoide, separó la ecuación de Laplace en coordenadas elipsoidales ϱ, μ y ν. Este proceso dio las mismas ecuaciones diferenciales ordinarias en cada una de las tres variables, a saber,

$$(\varrho^2-h^2)(\varrho^2-k^2)\frac{d^2E(\varrho)}{d\varrho^2} + \varrho(2\varrho^2-h^2-k^2)\frac{dE(\varrho)}{d\varrho} + \{(h^2+k^2)p - n(n+1)\varrho^2\}E(\varrho) = 0, \qquad (4)$$

[9] *Comm. Soc. Sci. Gott.*, 2, 1813 = *Werke*, 3, 123-162.
[10] *Werke*, 3, 207-230.
[11] *Jour. de Math.*, 2, 1837, 147-183; 4, 1839, 126-163.

con cambios apropiados para μ y ν en lugar de ϱ. Aquí h^2 y k^2 son los parámetros en las ecuaciones de las familias de superficies coordenadas y p y n son constantes. Esta ecuación es conocida como la ecuación diferencial de Lamé. Las soluciones $E(\varrho)$ son llamadas funciones de Lamé o armónicos elipsoidales. Para n entero, estas funciones caen dentro de cuatro clases de la forma

$$E_n^p(\varrho) = a_0\varrho^n + a_1\varrho^{n-2} + ...$$

o tales polinomios multiplicados por $\sqrt{p^2 - h^2}$ o $\sqrt{p^2 - k^2}$, o ambos factores. Para un valor dado de n, el número de tales funciones (para dar algunas propiedades de $E(\varrho)$) es $2n + 1$.

La segunda solución de la ecuación de Lamé (resultante de otras condiciones o propiedades de $E(\varrho)$) es

$$F_n^p(\varrho) = (2n + 1)E_n^p(\varrho) \int_\varrho^\infty \frac{d\varrho}{\sqrt{(\varrho^2 - h^2)(\varrho^2 - k^2)[E_n^p(\varrho)]^2}},$$

y tales funciones son llamadas funciones de Lamé de segunda clase. Estas fueron introducidas por Liouville [12] y Heine [13].

Las ecuaciones diferenciales

$$\frac{d^2\phi}{d\eta^2} + (a - 2k^2 \cos 2\eta)\phi = 0 \quad \frac{d^2\psi}{d\xi^2} - (a - 2k^2 \cosh 2\xi)\psi = 0$$

surgieron en el trabajo de Mathieu sobre las vibraciones de una membrana elíptica [14] y también surgieron en problemas sobre el potencial de cilindros elípticos cuando la separación de variables es aplicada a $\Delta u + k^2u = 0$ expresado en coordenadas cilíndricas elípticas en dos dimensiones. Estas coordenadas elípticas, incidentalmente, están relacionadas con las coordenadas rectangulares por las ecuaciones

$$x = h \cosh \xi \cos \eta, \qquad y = h \operatorname{senh} \xi \operatorname{sen} \eta,$$

donde $x = \pm h$, $y = 0$ son los focos de las elipses e hipérbolas cofocales de la familia de elipses y la familia de hipérbolas en el sistema

[12] *Jour. de Math.*, 10, 1845, 222-228.
[13] *Jour. für Math.*, 29, 1845, 185-208.
[14] *Jour. de Math.* (2), 13, 1868, 137-203.

de coordenadas elípticas. La variedad de formas en las que las ecuaciones de Mathieu están escritas, y las muchas notaciones para sus soluciones usadas por diferentes autores presenta una visión confusa. Las funciones definidas por cualquiera de las ecuaciones diferenciales fueron llamadas funciones de Heine del cilindro elíptico, y ahora son llamadas funciones de Mathieu. Mathieu y Heine fueron los primeros en obtener expresiones en series para las soluciones, y buscaron fijar el parámetro a de tal forma que una clase de soluciones sea periódica y de período 2π. El problema de encontrar soluciones periódicas —las más importantes para aplicaciones físicas— siguió a lo largo del siglo. En 1883 [15], Gaston Floquet (1847-1920) publicó una discusión completa de la existencia y propiedades de las soluciones periódicas de las ecuaciones diferenciales lineales de orden n-ésimo teniendo coeficientes periódicos con el mismo período w. Habiendo sido determinadas las propiedades generales de las soluciones, autores posteriores dedicaron considerable atención al problema de descubrir métodos prácticos para hallarlas. No se encontraron métodos generales (*cf.* sec. 7).

Heinrich Weber (1842-1913), en 1868, introdujo una clase de funciones especiales que se estudió extensamente [16]. Weber estaba interesado en integrar $\Delta u + k^2u = 0$ en un dominio encerrado por dos parábolas. Por tanto, el cambio de coordenadas rectangulares a coordenadas parabólicas (que son una clase límite de coordenadas elípticas) a través de la transformación

$$x = \xi^2 - \eta^2, \quad y = 2\xi\eta.$$

Para ξ = const., y para η = const., las dos familias de curvas son familias de parábolas, con cada miembro de una familia cortando los miembros de la otra ortogonalmente. Las ecuaciones diferenciales ordinarias que Weber derivó de la ecuación de ondas reducida por separación de variables son

$$\frac{d^2E}{d\xi^2} + (k^2\xi^2 + a)E = 0$$

$$\frac{d^2H}{d\eta^2} + (k^2\eta^2 - a)H = 0.$$

[15] *Ann de l'Ecole Norm. Sup.* (2), 12, 1883, 47-88.

[16] *Math. Ann.*, 1, 1869, 1-36.

Weber proporcionó cuatro soluciones particulares de la segunda ecuación en forma de integrales definidas. Las soluciones son llamadas funciones cilíndricas parabólicas. Weber demostró también que el único caso en el que la separación de variables puede ser aplicada a $\Delta u + k^2 u = 0$, entre todos los sistemas de coordenadas ortogonales, es el de superficies cofocales del segundo grado o casos particulares de él.

La clase de funciones especiales es aún más extensa de lo que podemos indicar aquí. Los tipos mencionados anteriormente, y muchos otros que fueron introducidos, sirven para resolver ecuaciones diferenciales en dominios acotados y representar funciones arbitrarias (usualmente las funciones iniciales de un problema de ecuaciones diferenciales parciales) en aquel dominio. La limitación a dominios acotados está impuesta por la propiedad de ortogonalidad. En el caso básico de las funciones trigonométricas este dominio es $(-\pi,\pi)$ porque, por ejemplo,

$$\int_{-\pi}^{\pi} \operatorname{sen} mx \operatorname{sen} nx \, dx = 0 \quad \text{en } m \neq n.$$

El problema de resolver ecuaciones diferenciales ordinarias sobre intervalos infinitos o intervalos semiinfinitos y de obtener expansiones de funciones arbitrarias sobre tales intervalos fue también atacado por muchos autores durante la segunda mitad del siglo, y funciones especiales como las funciones de Hermite, por primera vez introducidas por Hermite en 1864 [17], y Nikolai J. Sonine (1849-1915) en 1880 [18], sirvieron para resolver este problema.

Para trabajar con todos estos tipos de funciones especiales se debe conocer íntimamente sus propiedades de la misma manera que se conocen las propiedades de las funciones elementales. Como estas funciones especiales son más complicadas, las propiedades lo son de la misma manera. La literatura, tanto en artículos originales como en textos, es increíblemente vasta. Tratados completos han sido dedicados a las funciones de Bessel, funciones esféricas, funciones elipsoidales, funciones de Mathieu y de otros tipos.

---

[17] *Comp. Rend.*, 58, 1864, 93-100 y 266-273 = *Œuvres*, 2, 293-308.

[18] *Math. Ann.*, 16, 1880, 1-80.

3. La teoría de Sturm-Liouville

Los problemas implicando ecuaciones diferenciales parciales de la física matemática contienen comúnmente condiciones de frontera, tales como la condición de que la cuerda vibrante debe estar fija en los extremos. Cuando el método de separación de variables se aplica a una ecuación diferencial parcial, esta ecuación se descompone en dos o más ecuaciones diferenciales ordinarias, y las condiciones de frontera sobre la solución deseada se convierten en condiciones de frontera sobre una ecuación diferencial ordinaria. La ecuación ordinaria contiene generalmente un parámetro, que de hecho resulta del procedimiento de separación de variables y las soluciones se obtienen usualmente para valores particulares del parámetro. Estos valores son llamados los valores propios o valores característicos y la solución para cualquier valor propio es llamada una función propia. Más aún, para satisfacer la condicion o condiciones iniciales del problema original es necesario expresar una función $f(x)$ dada en términos de las funciones propias (véase, por ejemplo, [11] del cap. 28).

Estos problemas de determinar los valores propios y las funciones propias de una ecuación diferencial ordinaria con condiciones de frontera y de desarrollar una función dada en términos de una serie infinita de las funciones propias, que datan aproximadamente de 1750, se hicieron más prominentes al tiempo que se introducían nuevos sistemas de coordenadas y nuevas clases de funciones tales como las funciones de Bessel, los polinomios de Legendre, las funciones de Lamé y las funciones de Mathieu surgieron como funciones propias de ecuaciones diferenciales ordinarias. Dos autores, Charles Sturm (1803-1855), profesor de mecánica de la Sorbona, y Joseph Liouville (1809-1882), amigo de Sturm y profesor de matemáticas del Collège de France, decidieron atacar el problema general para cualquier ecuación diferencial de segundo orden. Sturm trabajó desde 1883 en problemas de ecuaciones diferenciales parciales, principalmente sobre el flujo de calor en una barra de densidad variable, y de aquí que fuera completamente consciente de los problemas de valores propios y funciones propias.

Las ideas matemáticas que aplicó a este problema [19] están estrechamente relacionadas con sus investigaciones de la «realidad» y distribución de las raíces de las ecuaciones algebraicas. Sus ideas en

[19] *Jour. de Math.*, 1, 1836, 106-186 y 373-444.

ecuaciones diferenciales, dice, provinieron de su estudio de ecuaciones en diferencias y de un paso al límite.

Liouville, informado por Sturm de los problemas sobre los que estaba trabajando, se dedicó a la misma materia [20]. Los resultados en los varios artículos de estos dos autores están bastante detallados, y son resumidos de manera más conveniente en notación moderna como sigue. Consideraron la ecuación general de segundo orden

$$Ly'' + My' + \lambda Ny = 0, \tag{5}$$

donde L, M y N son funciones continuas de x, L no es cero y λ es un parámetro. Tal tipo de ecuación puede ser transformada, multiplicando por $L^{-1}\, e^{\int ML^{-1}dx}$, en

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + \lambda\varrho(x)y = 0, \qquad p(x) > 0,\ \varrho(x) > 0.$$

Las condiciones en la frontera satisfechas por la ecuación original o transformada deben tener la forma general

$$\begin{aligned} y'(a) - h_1 y(a) &= 0, \\ y'(b) + h_2 y(b) &= 0 \end{aligned} \qquad h_1 \geq 0,\ h_2 \geq 0,\ a < b.$$

Sturm y Liouville demostraron los siguientes resultados fundamentales:

a) El problema tiene una solución que no es cero únicamente cuando λ toma cualquiera de los valores de la sucesión λ_n de números positivos que van hasta ∞.

b) Para cada λ_n, las soluciones son múltiplos de una función v_n, que se puede normalizar mediante la condición $\int_a^b \varrho v_n^2 dx = 1$.

c) Se mantiene la propiedad de ortogonalidad, $\int_a^b \varrho v_m v_n\, dx = 0$ para $m \neq n$.

d) Cada función f dos veces diferenciable en (a,b) y satisfaciendo las condiciones de frontera puede ser expandida en una serie uniformemente convergente.

$$f(x) = \sum_{n=1}^{\infty} c_n v_n(x)$$

---

[20] *Jour. de Math.*, 1, 1836, 253-265; 2, 1837, 16-35 y 418-436.

donde

$$c_n = \int_a^b \varrho f v_n(x)\, dx.$$

e) Se tiene la igualdad

$$\int_a^b pf^2\, dx = \sum_{n=1}^{\infty} c_n^2.$$

Esta última igualdad, llamada igualdad de Parseval, ya había sido demostrada formalmente por Marc-Antoine Parseval (¿-1836) en 1799 [21] para el conjunto de las funciones trigonométricas. De ahí se sigue la desigualdad demostrada por Bessel en 1828 [22], también para series trigonométricas, a saber,

$$\sum_{n=1}^{\infty} c_n^2 \leq \int_a^b |f(x)|^2\, dx.$$

De hecho, los resultados de Sturm-Liouville no fueron establecidos satisfactoriamente en todos sus aspectos. La demostración de que $f(x)$ puede ser representada como una suma infinita de las funciones propias fue inadecuada. Una dificultad era la cuestión de la completitud del conjunto de funciones propias, que para una $f(x)$ continua sobre (a,b) es la condición (d) anterior y que más o menos significa que el conjunto de las funciones propias es suficientemente grande como para representar «cualquier» $f(x)$. También la cuestión del sentido en el que la serie $\Sigma c_n v_n(x)$ converge a $f(x)$, ya fuera puntualmente, uniformemente o en algún sentido más general, no estaba precisado, aunque Liouville sí proporcionó demostraciones de convergencia en algunos casos, usando la teoría desarrollada por Cauchy y Dirichlet.

[21] *Mém. des sav. étrangers* (2), 1, 1805, 639-648.

[22] *Astronom. Nach.*, 6, 1828, 333-348.

4. Teoremas de existencia

Hemos notado con anterioridad, bajo este mismo tema de las ecuaciones diferenciales parciales del siglo XIX, que cuando los matemáticos encontraron más y más difícil el problema de obtener soluciones para ecuaciones diferenciales específicas, se hicieron la pregunta: ¿dada una ecuación diferencial, tiene una solución para condiciones iniciales y condiciones de frontera dadas? El mismo movimiento, igualmente esperado, ocurrió en las ecuaciones diferenciales ordinarias. Que la cuestión de la existencia fuese ignorada por tanto tiempo se debe en parte a que las ecuaciones diferenciales surgieron de problemas físicos y geométricos, y era intuitivamente claro que estas ecuaciones tenían soluciones.

Cauchy fue el primero en considerar la cuestion de la existencia de las soluciones de ecuaciones diferenciales y tuvo éxito al proporcionar dos métodos. El primero aplicable a

$$y' = f(x, y) \tag{6}$$

fue creado en algún momento entre 1820 y 1830 y resumido en sus *Exercices d'analyse* [23].

Este método, lo esencial del cual puede ser encontrado en Euler [24], utiliza la misma idea implicada en la integral como el límite de una suma. Cauchy deseaba demostrar que existe una y solamente una $y = f(x)$ que satisface (6) y que cumple la condición inicial de que $y_0 = f(x_0)$. Dividió (x_0, x) en n partes $\Delta x_0, \Delta x_1, \ldots, \Delta x_{n-1}$ y formó

$$y_{i+1} = y_i + f(x_i, y_i)\, \Delta x_i,$$

donde las x_i son cualquier valor de x en Δx_i. Entonces, por definición,

$$y_n = y_0 + \sum_{i=0}^{n-1} f(x_i, y_i)\, \Delta x_i.$$

Cauchy demuestra que cuando n se hace infinita, y_n converge a una única función

[23] Vol. 1, 1840, 327 sigs. = *Œuvres* (2), 11, 399-465.
[24] *Inst. Cal. Int.*, 1, 1768, 493.

$$y = y_0 + \int_{x_0}^{x} f(x, y)\, dx$$

y que esta función satisface (6) y las condiciones iniciales.

Cauchy supuso que $f(x,y)$ y f_y son continuas para todos los valores reales de x e y en el rectángulo determinado por los intervalos (x_0,x) e (y_0,y). En 1876, Rudolph Lipschitz (1832-1903) debilitó las hipótesis del teorema [25]. Su condición esencial fue que para todas las (x,y_1) y (x,y_2) en el rectángulo $|x - x_0| \leqslant a$, $|y - y_0| \leqslant b$, esto es, para dos puntos cualesquiera con la misma abcisa, existe una constante k tal que

$$|f(x, y_1) - f(x, y_2)| < K(y_1 - y_2).$$

Esta condición es conocida como la condición de Lipschitz y el teorema de existencia es llamado teorema de Cauchy-Lipschitz.

El segundo método de Cauchy para establecer la existencia de soluciones para ecuaciones diferenciales, el método de las funciones dominantes o mayorantes es más extensamente aplicable que el primero, y fue aplicado por Cauchy en el dominio complejo. El método fue presentado en una serie de ensayos en las *Comptes Rendus* durante los años de 1839 a 1842 [26]. El método fue llamado por Cauchy *calcul des limites* porque proporciona límites inferiores dentro de los cuales la solución cuya existencia se establece converge seguramente. Briot y Bouquet simplificaron el método y su versión [27] se convirtió en la habitual.

Para ilustrar el método veamos cómo se aplica a

$$y' = f(x, y),$$

donde f es analítica en x e y. El teorema que ha de ser establecido dice así: si para

$$\frac{dy}{dx} = f(x, y) \tag{7}$$

[25] *Bull. des Sci. Math.* (1), 10, 1876, 149-59.

[26] *Œuvres* (1), vols. 4 a 7 y 10. Los artículos más importantes están en las *Comptes Rendus* de 5 de agosto y 21 de noviembre de 1839; 29 de junio, 26 de octubre, 2 de noviembre y 9 de noviembre de 1840; y 20 de junio y 4 de julio de 1842.

[27] *Comp. Rend.*, 39, 1854, 368-371.

la función $f(x,y)$ es analítica en un entorno de $P_0 = (x_0,y_0)$, la ecuación diferencial tiene entonces una solución única $y(x)$ que es analítica en el entorno de x_0 y que se reduce a y_0 cuando $x = x_0$. La solución puede ser representada por la serie

$$y = y_0 + y_0'(x - x_0) + \frac{y_0''}{2!}(x - x_0)^2 + \frac{y_0''}{3!}(x - x_0)^3 + \ldots \qquad (8)$$

donde $y_0' = dy/dx$ en (x_0,y_0) y análogamente para y_0'', y_0''',..., y donde las derivadas están determinadas mediante la diferenciación sucesiva de la ecuación diferencial original en la que y es tratada como función de x.

El método de demostración, que esbozamos solamente, se vale en primera instancia del hecho de que $f(x,y)$ es analítica en un entorno de (x_0,y_0), que por conveniencia tomamos como $(0,0)$: existe un círculo de radio a alrededor de $x_0 = 0$ y un círculo de radio b alrededor de $y_0 = 0$ en el cual $f(x,y)$ es analítica. Entonces $f(x,y)$ tiene una cota superior M para todos los valores de x e y en los círculos respectivos. Ahora, el propio método para obtener la serie (8) garantiza que satisface formalmente (7). El problema es demostrar que la serie converge.

Para este fin se introduce la función mayorante

$$F(x, y) = \sum \frac{M}{a^p b^q} x^p y^q,$$

que es el desarrollo de

$$F(x, y) = \frac{M}{(1 - x/a)(1 - y/b)}. \qquad (9)$$

Luego se demuestra que la serie solución de

$$\frac{dY}{dx} = F(x, Y) \qquad (10)$$

a saber,

$$Y = Y_0'x + Y_0''\frac{x^2}{2!} + Y_0''\frac{x^3}{3!} + \ldots, \qquad (11)$$

que es derivada de (10) de la misma manera que (8) es derivada de (7), domina término a término la serie (8). De aquí que si (11) converge, entonces lo hace (8). Para demostrar que (11) converge debe resolverse (10) usando explícitamente el valor de F en (9) demostrando que la expansión de la serie de la solución, que ha de ser (11), converge.

El método, por sí mismo, no determina el radio preciso de convergencia de la serie para y. Numerosos esfuerzos se dedicaron, por tanto, a demostrar que el radio puede ser extendido. Sin embargo, los artículos no proporcionan el dominio total de convergencia y son de poca importancia práctica.

Liouville publica por primera vez un tercer método para establecer la existencia, para una ecuación de segundo orden, de soluciones de ecuaciones diferenciales ordinarias, el cual probablemente era conocido por Cauchy [28]. Este es el método de aproximaciones sucesivas y ahora se le atribuye a Emile Picard, dado que proporcionó el método en la forma general [29]. Para la ecuación en x e y reales,

$$y' = f(x, y),$$

donde $f(x,y)$ es analítica en x e y, y cuya solución $y = f(x)$ debe pasar por (x_0,y_0) el método es introducir la sucesión de funciones

$$y_1(x) = y_0 + \int_{x_0}^{x} f(t, y_0)\, dt$$

$$y_2(x) = y_0 + \int_{x_0}^{x} f(t, y_1(t))\, dt$$

....

$$y_n(x) = y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t))\, dt.$$

Entonces se demuestra que $y_n(x)$ tiende a un límite $y(x)$ que es la única función continua de x que satisface la ecuación diferencial

[28] *Jour. de Math.* (1), 3, 1838, 561-614.
[29] *Jour. de Math.* (4), 6, 1890, 145-210; y (4), 9, 1893, 217-271.

ordinaria y tal que $y(x_0) = y_0$. El método, como comúnmente se le presenta hoy día, presupone que $F(x,y)$ sólo satisface la condición de Lipschtiz. El método fue extendido a ecuaciones de segundo orden por Picard en el artículo de 1893 y también se ha extendido a x e y complejos.

Los diversos métodos descritos anteriormente no sólo se aplicaron a ecuaciones diferenciales ordinarias de orden superior, sino también a sistemas de ecuaciones diferenciales para variables con valores complejos. De esta manera, Cauchy extendió su segundo tipo de teoremas de existencia a sistemas de ecuaciones diferenciales ordinarias de primer orden en n variables dependientes. También, con su método del *calcul des limites,* trató sistemas en el dominio complejo [30]. El resultado de Cauchy es como sigue: dado el sistema de ecuaciones

$$\frac{dy_k}{dx} = f_k(x, y_0, ..., y_{n-1}), \qquad k = 0, 1, 2, ..., n - 1, \qquad (12)$$

sean $f_0,...,f_{n-1}$, funciones monógenas (analíticas univalentes) de sus argumentos y supongamos que sean desarrollables en un entorno de los valores iniciales

$$x = \xi, y_0 = \eta_0, ..., y_{n-1} = \eta_{n-1}$$

en potencias enteras positivas de

$$x - \xi, y_0 - \eta_0, ..., y_{n-1} - \eta_{n-1}.$$

Entonces, existen n series de potencias en $x - \xi$, convergentes en el entorno de $x = \xi$ las cuales, cuando son sustituidas en las $y_0, ..., y_{n-1}$ en (12), satisfacen las ecuaciones. Estas series de potencias son únicas, proporcionan una solución regular del sistema y toman los valores iniciales. En esta generalidad, el resultado puede encontrarse en la «Mémoire sur l'emploi du nouveau calcul, appelé calcul des limites, dans l'intégration d'un système d'équations différentielles» («Memoria sobre el empleo del nuevo cálculo, llamado cálculo de

[30] Para diferentes demostraciones de existencia para sistemas de ecuaciones diferenciales de primer orden, véase la segunda referencia de Painlevé en la bibliografía al final del capítulo.

límites, en la integración de un sistema de ecuaciones diferenciales»)[31]. La idea era darse satisfacción con el establecimiento de la existencia y con obtener la solución en el entorno de un punto en el plano complejo. Weierstrass obtuvo igual resultado en el mismo año (1842), pero no lo publicó hasta que aparecieron sus *Werke (Obras)* en 1894[32].

5. La teoría de singularidades

A mediados del siglo XIX, el estudio de las ecuaciones diferenciales ordinarias tomó un nuevo curso. Los teoremas de existencia y la teoría de Sturm-Liouville presuponen que las ecuaciones diferenciales contienen funciones analíticas, o, al menos, funciones continuas en el dominio donde las soluciones son consideradas. Por otro lado, algunas de las ecuaciones diferenciales ya consideradas, tales como las de Bessel, Legendre y la ecuación hipergeométrica, cuando son expresadas de manera que el coeficiente de la segunda derivada sea la unidad, tienen coeficientes que son singulares, y la forma de las soluciones en serie en los entornos de los puntos singulares, particularmente los de la segunda solución, es peculiar. A partir de aquí los matemáticos se dirigieron al estudio de las soluciones en los entornos de los puntos singulares, esto es, puntos en los cuales uno o más de los coeficientes son singulares. El punto en el que todos los coeficientes son al menos continuos y (comúnmente) analíticos es llamado punto ordinario.

Las soluciones en los entornos de puntos singulares son obtenidas como series y el conocimiento de la forma adecuada de las series debe estar disponible antes de calcularlas. Este conocimiento sólo se obtiene a partir de la ecuación diferencial. El nuevo problema fue descrito por Lazarus Fuchs (1833-1902) en un artículo de 1866 (véase más abajo). «En la situación actual de la ciencia el problema de la teoría de las ecuaciones diferenciales no es tanto reducir una ecuacion dada a cuadraturas, como deducir a partir de la misma ecuación el comportamiento de sus integrales en todos los puntos del plano, esto es, para todos los valores de la variable compleja.» Para este problema, fue el trabajo de Gauss sobre series hipergeométricas el

[31] *Comp. Rend.*, 15, 1842, 14-25 = *Œuvres* (1), 7, 5-17.

[32] *Meth. Werke*, 1, 75-85.

que señaló el camino. Las grandes figuras fueron aquí Riemann y Fuchs, este último alumno de Weierstrass y su sucesor en Berlín. La teoría resultante es llamada teoría fuchsiana de las ecuaciones diferenciales lineales.

La atención en esta nueva área se concentró sobre las ecuaciones diferenciales lineales de la forma

$$y^{(n)} + p_1(z)y^{(n-1)} + \ldots + p_n(z)y = 0, \qquad (13)$$

donde las $p_i(z)$ son funciones analíticas univalentes de z compleja, excepto en puntos singulares aislados. Esta ecuación fue muy estudiada porque sus soluciones abarcaban todas las funciones elementales y aun algunas funciones superiores, tales como las funciones modulares y automorfas que encontraremos más adelante.

Antes de considerar las soluciones en y dentro de un entorno de puntos singulares, señalemos un teorema básico que se sigue del teorema de existencia de Cauchy para sistemas de ecuaciones diferenciales ordinarias, probado directamente por Fuchs [33], aunque reconoció su deuda con las clases de Cauchy. Si los coeficientes $p_1, \ldots, p_n$ son analíticos en el punto a y en algún entorno de ese punto, y si están dadas condiciones iniciales arbitrarias para y y sus primeras $n - 1$ derivadas en $z = a$, entonces existe una solución en serie de potencias única para y en términos de z de la forma

$$y(z) = \sum_{r=0}^{\infty} \frac{1}{r!} y^{(r)}(a)(z - a)^r. \qquad (14)$$

Fuchs añadió al resultado de Cauchy que la serie es absoluta y uniformemente convergente dentro de cualquier círculo teniendo a a como centro y en el cual las $p_i(z)$ son analíticas. Se sigue que las soluciones pueden poseer singularidades únicamente donde los coeficientes son singulares.

Briot y Bouquet iniciaron el estudio de las soluciones en los entornos de los puntos singulares [34]. Ya que sus resultados para ecuaciones lineales de primer orden fueron rápidamente generalizados, consideraremos los tratamientos más generales.

A fin de llegar al comportamiento en el entorno de los puntos

[33] *Jour. für Math.*, 66, 1866, 121-160 = *Math. Werke*, 1, 159 sigs.
[34] *Jour. d'Ecole Poly.* (1), 21, 1856, 85-132, 133-198, 199-254.

singulares, Riemann propuso un enfoque poco usual. A pesar de que las $p_i(z)$ en (13) se suponen funciones analíticas univalentes, excepto en puntos singulares aislados, las soluciones $y_i(z)$, analíticas excepto posiblemente en los puntos singulares, no son en general univalentes sobre el dominio entero de los valores de z. Supongamos que tenemos un conjunto fundamental de soluciones, $y_i(z)$, $i = 1, 2, ..., n$, esto es, n soluciones independientes de la clase especificada en el teorema anterior. Entonces la solución general es

$$y = c_1y_1 + c_2y_2 + ... + c_ny_n,$$

donde las c_i son constantes.

Si ahora rastreamos el comportamiento de una y_i analítica a lo largo de una trayectoria cerrada envolviendo un punto singular, y_i cambiará entonces su valor a otra rama de la misma función aunque permanezca no obstante como solución de la ecuación diferencial. Como cualquier solución es una combinación lineal de n soluciones particulares, las nuevas y_i, digamos y_i', son una combinación de las y_i. Así obtenemos

$$\begin{aligned} y_1' &= c_{11}y_1 + ... + c_{1n}y_n \\ y_2' &= c_{21}y_1 + ... + c_{2n}y_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n' &= c_{n1}y_1 + ... + c_{nn}y_n. \end{aligned} \tag{15}$$

Esto es, las $y_1, ..., y_n$ sufren una cierta transformación lineal cuando cada una es llevada alrededor de una trayectoria cerrada rodeando un punto singular. Tal transformación surge para cualquier trayectoria alrededor de los puntos circulares o combinación de los puntos singulares. El conjunto de las transformaciones forma un grupo [35], llamado el grupo de monodromía de la ecuación diferencial, término que introdujo Hermite [36].

El enfoque de Riemann para obtener el carácter de las soluciones

---

[35] En la época de Riemann era conocida la noción algebraica de grupo. En este libro será introducida en el cap. 31. Sin embargo, lo único que se necesita conocer aquí es que la aplicación de dos transformaciones sucesivas es una transformación del conjunto y que la inversa de cada transformación pertenece al conjunto.

[36] *Comp. Rend.*, 32, 1851, 458-461 = *Œuvres*, 1, 276-280.

en el entorno de los puntos singulares apareció en su artículo de 1857 «Beitrage zur Theorie die Gauss'sche Reihe $F(\alpha,\beta,\gamma,z)$ darstellbaren Functionen» [37] («Contribución a la teoría gaussiana de las series $F(\alpha,\beta,\gamma,x)$ representando funciones»). La ecuación diferencial hipergeométrica, como Gauss la llamaba, tiene tres puntos singulares, 0, 1 y ∞. Riemann demostró que, para x compleja, para obtener conclusiones acerca del comportamiento de las soluciones particulares alrededor de los puntos singulares de la ecuación de segundo orden, no es necesario conocer la propia ecuación diferencial, sino más bien saber cómo dos soluciones independientes se comportan cuando la variable independiente recorre trayectorias cerradas alrededor de tres puntos singulares. Esto es, debemos conocer las transformaciones

$$y_1' = c_{11}y_1 + c_{12}y_2, \quad y_2' = c_{21}y + c_{22}y_2$$

para cada punto singúlar.

De esta manera, la idea de Riemann, al tratar funciones definidas por ecuaciones diferenciales, fue derivar las propiedades de las funciones de un conocimiento del grupo de monodromía. Su artículo de 1857 trataba la ecuación diferencial hipergeométrica, pero se planteaba tratar las ecuaciones diferenciales ordinarias lineales de orden n-ésimo con coeficientes algebraicos. En un fragmento escrito en 1857, no publicado hasta que sus obras escogidas aparecieron en 1876 [38], Riemann consideró ecuaciones más generales que las de segundo orden con tres puntos singulares. Consecuentemente, supone que tiene n funciones uniformes, finitas y continuas, excepto en ciertos puntos asignados arbitrariamente (los puntos singulares) y sometidos a una substitución lineal fijada arbitrariamente cuando z describe un circuito cerrado alrededor de cierto punto. Demuestra entonces que tal sistema de funciones satisface una ecuación diferencial lineal de orden n-ésimo. Pero no demuestra que los puntos de ramificación (puntos singulares) y las sustituciones puedan ser escogidas arbitrariamente. Aquí su trabajo fue incompleto y dejó abierto un problema conocido como el problema de Riemann: dados m puntos $a_1, \ldots, a_m$ en el plano complejo, y asociada a cada uno de ellos una transformación lineal de la forma (15), demostrar sobre la base

[37] *Werke*, 67-83.
[38] *Werke*, 379-390.

de suposiciones elementales acerca del comportamiento del grupo de monodromía asociado con estos puntos singulares (siempre y cuando tal comportamiento no esté ya determinado) que está determinada una clase de funciones $y_1,..., y_n$ que satisface una ecuación diferencial lineal de orden n-ésimo con las a_i dadas como puntos singulares y tales que cuando la z recorre una trayectoria cerrada alrededor de las a_i, las y_i sufren la transformación lineal asociada con las a_i.

Guiado por el ensayo de Riemann de 1857 sobre la ecuación hipergeométrica, Fuchs llevó más lejos el trabajo sobre singularidades. Fuchs y sus estudiantes, empezando en 1865 [39], estudiaron las ecuaciones diferenciales de orden n-ésimo, mientras que Riemann habia publicado nada más que sobre la ecuación diferencial hipergoemétrica de Gauss. Fuchs no siguió el enfoque de Riemann, sino que trabajó directamente con la ecuación diferencial lineal, hasta extender totalidad de la teoría de ecuaciones diferenciales al dominio de la teoría de funciones complejas.

En los ensayos mencionados anteriormente, Fuchs publicó su trabajo principal sobre ecuaciones diferenciales ordinarias. Empieza con la ecuación diferencial lineal de orden n cuyos coeficientes son funciones racionales de x. Mediante un examen cuidadoso de la convergencia de las series que satisfacen formalmente la ecuación, encuentra que los puntos singulares de la ecuación son fijos, esto es, independientes de las constantes de integración, y pueden ser encontrados antes de integrar, ya que son los polos de los coeficientes de la ecuación diferencial.

Más adelante demuestra que un sistema fundamental de soluciones sufre una transformación lineal con coeficientes constantes cuando la variable independiente z describe un circuito encerrando un punto singular. De este comportamiento de las soluciones deriva expresiones válidas para ellas en una región circular encerrando aquel punto y extendiéndose al punto singular siguiente. Establece la existencia de sistemas de n funciones uniformes, finitas y continuas con excepción de los entornos de ciertos puntos y sometidas a sustituciones lineales con coeficientes constantes cuando la variable z describe circuitos cerrados alrededor de estos puntos.

Luego Fuchs consideró qué propiedades debe tener una ecuación

[39] *Jour. für Math.*, 66, 1866, 121-160; 68, 1868, 354-385.

diferencial de la forma (13) en función que sus soluciones en un punto singular $z = a$ tengan la forma

$$(z - a)^s[\phi_0 + \phi_1 \log (z - a) + ... + \phi_\lambda \log^\lambda (z - a)],$$

donde s es algún número (que más adelante puede ser especificado) y las ϕ_i pueden ser funciones univalentes en un entorno de $z = a$ que pueden tener polos de orden finito. Su respuesta fue que una condición necesaria y suficiente es que $p_r(z) = (z - a)^{-r} P(z)$, donde $P(z)$ es analítica cerca de $z = a$. Así, $p_1(z)$ tiene un polo de orden uno y, así sucesivamente. Tal punto a es llamado punto singular regular (Fuchs lo llamó un punto de determinación).

Fuchs estudió también una clase más específica de ecuaciones de la forma (13). Una ecuación lineal homogénea de este tipo es llamada de *tipo fuchsiano* cuando tiene en el peor de los casos puntos singulares regulares en el plano complejo extendido (incluyendo el punto del ∞). En este caso las $p_i(z)$ deben ser funciones racionales de z. Por ejemplo, la ecuación hipergeométrica tiene puntos singulares regulares en $z = 0$, 1 y ∞.

Pero el estudio de las integrales de las ecuaciones diferenciales en un entorno de un punto dado no proporciona necesariamente las propias integrales. El estudio fue tomado como el punto de partida para la investigación de las integrales completas. A partir de las grandes investigaciones de Fuchs, los matemáticos han tenido éxito en extender la clase de ecuaciones diferenciales ordinarias lineales que pueden ser integradas explícitamente. Con anterioridad, únicamente las ecuaciones lineales de orden n con coeficientes constantes y la ecuación de Legendre

$$(ax+b)^n \frac{d^n y}{dx^n}+A(ax+b)^{n-1} \frac{d^{n-1} y}{dx^{n-1}}+...+L(ax+b) \frac{dy}{dx}+My=0$$

podían ser integradas, la última mediante la transformación $ax + b = e^t$. Las nuevas ecuaciones susceptibles de ser integradas son las que contienen integrales que son funciones uniformes (univalentes) de z. Se reconoce que las integrales tienen esta propiedad mediante el estudio de los puntos singulares de la ecuación diferencial. Las integrales generales así obtenidas, por regla general, son funciones nuevas.

Más allá de resultados generales sobre los tipos de integrales que

pueden tener clases especiales de ecuaciones diferenciales, existe una aproximación con series a las soluciones en el punto $z = a$ donde la ecuación tiene un punto singular regular. Si el origen es el punto, entonces la ecuación tiene la forma

$$z^n \frac{d^n w}{dz^n} + z^{n-1} P_1(z) \frac{d^{n-1} w}{dz^{n-1}} + \dots + z P_{n-1}(z) \frac{dw}{dz} + P_n(z) w = 0,$$

en las que $P_i(z)$ son analíticas en y alrededor de $z = 0$. En este caso se puede obtener las n soluciones fundamentales en la forma de series cerca de $z = 0$ y demostrar que la serie converge para algún rango de valores de z. Las series son de la forma

$$w = \sum_{v=0}^{\infty} c_v z^{\varrho + v}$$

y las ϱ y c_v son determinables para cada solución. El resultado se debe a Georg Frobenius (1849-1917) [40].

El problema de Riemann también fue atacado durante la última parte del siglo XIX, pero sin éxito hasta que Hilbert en 1905 [41] y Oliver D. Kellogg (1878-1932) [42], con la ayuda de la teoría de ecuaciones integrales, que había sido desarrollada mientras tanto, proporcionaron la primera solución completa. Demostraron que la transformación generadora del grupo de monodromía puede ser prescrita arbitrariamente.

6. Las funciones automorfas

Poincaré y Felix Klein estudiaron a continuación la teoría de las soluciones de las ecuaciones diferenciales lineales. La teoría que ellos introdujeron es llamada funciones automorfas, materia que, aunque importante por otras varias aplicaciones, juega un papel fundamental en la teoría de ecuaciones diferenciales.

---

[40] *Jour. für Math.*, 76, 1874, 214-235 = *Ges. Abh.*, 1, 84-105.

[41] *Proc. Third Internat. Math. Cong.*, 1905, 233-240; y *Nachrichten König. Ges. der Wiss. zu Gött.*, 1905, 307-388. También en D. Hilbert, *Grudzüge einer allgemeinen Theorie der linearen Integralgleichungen*, 1912, Chelsea (reimp.), 1953, 81-108.

[42] *Math. Ann.*, 60, 1905, 424-433.

Henri Poincaré (1854-1912) era profesor de la Sorbona. Sus publicaciones, casi tan numerosas como las de Euler y Cauchy, cubren un rango muy amplio de matemáticas y fisica matemática. Sus investigaciones en física, que no tendremos ocasión de discutir, incluyeron la atracción capilar, elasticidad, teoría del potencial, hidrodinámica, propagación del calor, electricidad, óptica, teoría electromagnética, relatividad y sobre todo, mecánica celeste. Poincaré tenía un visión penetrante y de cada problema que estudió obtuvo su carácter esencial. Enfocaba agudamente un problema y lo examinaba detalladamente. También creía en el estudio cualitativo de todos los aspectos del problema.

Las funciones automorfas son generalizaciones de las funciones circulares, hiperbólicas, elípticas u otras funciones del análisis elemental. La funcion *sen* z no cambia de valor si z es reemplazada por $z + 2m\pi$, donde m es cualquier entero. También se puede decir que la función se mantiene sin alteración en su valor si z está sujeta a cualquier transformación del grupo $z' = z + 2m\pi$. La función hiperbólica *senh* z se mantiene sin alteración en su valor si z está sujeta a cualquier transformación del grupo $z' = z + 2\pi mi$. Una función elíptica se mantiene invariante en su valor bajo las transformaciones del grupo $z' = z + mw + m'w'$ donde w y w' son los períodos de la función. Todos estos grupos son discontinuos (término introducido por Poincaré); esto es, todas las transformaciones de cualquier punto bajo las transformaciones del grupo son finitas en número en cualquier dominio acotado cerrado.

El término función automorfa es usado ahora para incluir funciones que son invariantes bajo el grupo de transformaciones

$$z' = \frac{az + b}{cz + d}, \tag{16}$$

donde a, b, c y d puede ser números reales o complejos, y $ad - bc = 1$, o bajo algún subgrupo de este grupo. Más aún, el grupo debe ser discontinuo en cualquier parte finita del plano complejo.

Las primeras funciones automorfas estudiadas fueron las funciones modulares elípticas. Estas funciones son invariantes bajo el grupo modular, que es aquel subgrupo de (16) donde a, b, c y d son enteros reales y $ad - bc = 1$, o bajo algún subgrupo de este grupo. Estas funciones modulares elípticas se derivan de las funciones elíp-

ticas. No las seguiremos aquí, ya que no entran en la teoría básica de las ecuaciones diferenciales.

Las funciones automorfas más generales fueron introducidas para estudiar ecuaciones diferenciales lineales de segundo orden

$$\frac{d^2\eta}{dz^2} + p_1 \frac{d\eta}{dz} + p_2\eta = 0, \tag{17}$$

donde p_1 y p_2 fueron en principio funciones racionales de z. Un caso especial es la ecuación hipergoemétrica

$$\frac{d^2\eta}{dz^2} + \frac{\gamma - (\alpha + \beta + 1)}{z(1 - z)} \frac{d\eta}{dz} + \frac{\alpha\beta}{z(z - 1)} \eta = 0 \tag{18}$$

con los tres puntos singulares 0, 1 e ∞.

Riemann, en sus clases de 1858-1859 sobre las series hipergeométricas y en un ensayo póstumo de 1867 sobre superficies mínimas, e, independientemente, Schwarz [43] establecieron lo siguiente. Sean η_1 y η_2 dos soluciones particulares cualesquiera de la ecuación (17). Todas las soluciones son expresadas como

$$\eta = m\eta_1 + n\eta_2.$$

Cuando z recorre una trayectoria cerrada alrededor de un punto singular, η_1 y η_2 pasan a ser

$$\eta_1^1 = a\eta_1 + b\eta_2, \qquad \eta_2^1 = c\eta_1 + d\eta_2$$

y permitiendo que z recorra trayectorias cerradas alrededor de todos los puntos singulares se obtiene el grupo completo de tales transformaciones lineales, que es el grupo de monodromía de la ecuación diferencial.

Ahora, sea $\zeta(z) = \eta_1/\eta_2$. Cuando z recorre una trayectoria cerrada, el cociente ζ es transformado en

$$\zeta^1 = \frac{a\zeta + b}{c\zeta + d}. \tag{19}$$

[43] *Jour. für Math.*, 75, 1873, 292-335 = *Ges. Abh.*, 2, 211-259.

A partir de (17) encontramos que ζ satisface la ecuación diferencial

$$\frac{\zeta''}{\zeta'} - \frac{3}{2}\cdot\left(\frac{\zeta''}{\zeta'}\right) = 2p_2 - \frac{1}{2}p_1^2 - p_1'. \tag{20}$$

Si tomamos para las p_1 y p_2 en (17) las funciones particulares de (18) obtenemos

$$\frac{\zeta''}{\zeta'} - \frac{3}{2}\left(\frac{\zeta''}{\zeta'}\right)^2 = \frac{1-\lambda^2}{2z^2} + \frac{1-\mu^2}{2(1-z)^2} - \frac{\lambda^2+\mu^2-\nu^2-1}{2z(1-z)}, \tag{21}$$

donde $\lambda^2 = 1 - \gamma^2$, $\mu^2 = (\gamma - \alpha - \beta)^2$, $\gamma^2 = (\alpha - \beta)^2$ y λ, μ, ν son tomadas positivas (α, β, γ son reales). La clase de transformaciones (19) es el grupo de monodromía de la ecuación diferencial (21).

Después Riemann y Schwarz demostraron que toda solución particular $\zeta(z)$ de la ecuación (21), cuando λ, μ y ν son reales, es una aplicación conforme de la parte superior del plano z (Fig. 29.1) dentro de un triángulo curvilíneo con arcos circulares en el plano ζ cuyos ángulos son $\lambda\pi$, $\mu\pi$ y $\nu\pi$.

En el caso de un domino acotado por tres arcos de círculos, si los ángulos del triángulo satisfacen ciertas condiciones, la función inversa de $\zeta = \zeta(z)$ es una función automorfa $z = \phi(\zeta)$ cuyo dominio completo de existencia es un semiplano o un círculo. Esta función permanece invariante bajo la transformación de ζ por elementos del grupo de transformaciones lineales (19), que llevan cualquier triángulo curvilíneo de la forma mostrada en otro. El triángulo «circular» dado es el dominio fundamental de grupo. Bajo el grupo de transformaciones, este dominio es aplicado en triángulos análogos cuya unión cubre el semiplano o el círculo. El triángulo circular es análogo al paralelogramo en el caso de las funciones elípticas.

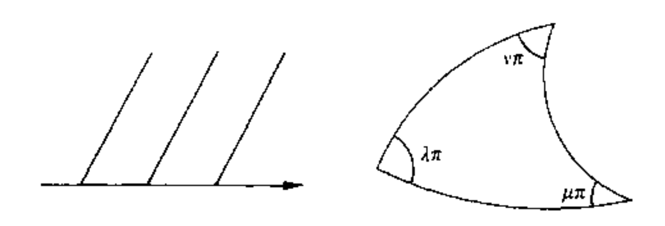


FIGURA 29.1

El trabajo de Poincaré y Klein arranca de este punto. Klein realizó cierto trabajo básico sobre funciones automorfas antes de 1880. Durante los años 1881-1882, trabajó con Poincaré, quien también había llevado a cabo algún trabajo previo sobre la materia después de que su atención fue atraída por el trabajo de Fuchs mencionado con anterioridad. En 1884, Poincaré publicó cinco artículos importantes sobre funciones automorfas en los primeros cinco volúmenes del *Acta Mathematica.* Cuando el primero de éstos fue publicado en el primer volumen de la nueva *Acta Mathematica,* Kronecker advirtió al editor Mittag-Leffler que este artículo inmaduro y obscuro mataría a la revista.

Guiado por la teoría de funciones elípticas, Poincaré inventó una nueva clase de funciones automorfas [44]. Esta clase fue obtenida considerando la función inversa del cociente de dos soluciones linealmente independientes de la ecuación

$$\frac{d^2\eta}{dz^2} + P(w, z)\frac{d\eta}{dz} + Q(w, z)\eta = 0,$$

donde w y z están relacionadas por una ecuación polinomial $\phi(w,z) = 0$, y P y Q son funciones racionales. Esta es la clase de las funciones automorfas fuschsianas y consiste en funciones (univalentes) meromorfas dentro de un círculo (llamado círculo fundamental) que son invariantes bajo la clase de transformaciones lineales de la forma

$$z' = \frac{az + b}{cz + d}, \tag{22}$$

donde a, b, c y d son reales y $ad - bc = 1$. Estas transformaciones que dejan el círculo y su interior invariantes forman un grupo llamado grupo fuchsiano. La función de Schwarz $\phi(\zeta)$ constituye el ejemplo más sencillo de una función fuchsiana. Poincaré demostró la existencia de una clase de funciones automorfas más general que las funciones elípticas modulares [45].

---

[44] *Acta Math.*, 1, 1882, 1-62 y 193-294 = *Œuvres,* 2, 108-168, 169-257.

[45] En este trabajo sobre grupos fuchsianos, Poincaré utilizó la geometría no euclídea (Cap. 36) y demostró que el estudio de los grupos fuchsianos se reduce al del grupo de traslaciones de la geometría de Lobatchevski.

La construcción de Poincaré de las funciones automorfas (en el segundo artículo de 1882) estaba basada en su serie theta. Sean las transformaciones del grupo (22)

$$z' = \frac{a_i z + b_i}{c_i z + d_i}, \quad a_i d_i - b_i c_i = 1, \quad i = 1, 2, \ldots \tag{23}$$

Sean z_1, z_2,... las transformadas de z por las diversas transformaciones del grupo. Sea $H(z)$ una función racional (al margen de otras consideraciones menores). Entonces, la serie theta de Poincaré es la función

$$\theta(z) = \sum_{i=0}^{\infty} (c_i z + d_i)^{-2m} H(z_i), \quad m > 1. \tag{24}$$

Se puede demostrar que $\theta(z_j) = (c_j z + d_j)^{2m}\ \theta(z)$. Ahora, sean $\theta_1(z)$ y $\theta_2(z)$ las dos series theta con la misma m. Estas series no son sólo funciones uniformes sino enteras. Entonces

$$F(z) = \frac{\theta_1(z)}{\theta_2(z)} \tag{25}$$

es una función automorfa del grupo (23). Poincaré llamó la serie (24) serie theta-fuchsiana o serie theta-kleiniana según que el grupo al cual pertenece sea fuchsiano o kleiniano (este último será descrito dentro de un momento).

Las funciones fuchsianas son de dos tipos, uno que existe sobre todo el plano, y otro que sólo existe en el interior del círculo fundamental. La función inversa de una función fuchsiana es, como vimos anteriormente, el cociente de dos integrales de una ecuación diferencial lineal de segundo orden con coeficientes algebraicos. Tal ecuación, que Poincaré llamó ecuación fuchsiana, puede ser integrada por medio de funciones fuchsianas.

Más adelante, Poincaré [46] extendió el grupo de transformaciones (22) a coeficientes complejos y consideró varios tipos de tales grupos, a los que llamó kleinianos. Nos debemos conformar aquí con notar que un grupo es kleiniano si, esencialmente, no es finito ni fuchsiano sino que, por supuesto, es de la forma (22) y discontinuo

[46] *Acta Math.*, 3, 1883, 49-92; 4, 1884, 201-312 = *Œuvres*, 2, 258-299, 300-401.

en cualquier parte del plano complejo. Para estos grupos kleinianos, Poincaré obtuvo nuevas funciones automorfas, esto es, funciones invariantes bajo los grupos kleinianos, a las que llamó funciones kleinianas. Estas funciones tienen propiedades análogas a las fuchsianas; sin embargo, la región fundamental para las nuevas funciones es más complicada que un círculo. Incidentalmente, Klein había considerado las funciones fuchsianas, mientras que Lazarus Fuchs no. Klein, por tanto, protestó contra Poincaré. Poincaré respondió llamando a la siguiente clase de funciones automorfas, a pesar de que él mismo las descubrió, kleinianas, ya que —como alguien observó perversamente— éstas nunca fueron consideradas por Klein.

Más adelante, Poincaré mostró cómo expresar las integrales de ecuaciones lineales de orden n-ésimo con coeficientes *algebraicos*, teniendo únicamente puntos singulares regulares, con ayuda de las funciones kleinianas. Así, esta clase completa de ecuaciones diferenciales lineales se resuelve mediante el uso de las nuevas funciones trascendentes de Poincaré.

7. El trabajo de Hill sobre soluciones periódicas de las ecuaciones lineales

Mientras que estaba siendo creada la teoría de funciones automorfas, el trabajo en astronomía estimuló el interés sobre una ecuación diferencial ordinaria de segundo orden, que de alguna manera era más general que la ecuación de Mathieu. Toda vez que el problema de los n cuerpos no estaba resuelto explícitamente y sólo se disponía de soluciones en serie muy complicadas, los matemáticos volvieron a preferir soluciones periódicas.

La importancia de las soluciones periódicas surge del problema de la estabilidad de una órbita de un planeta o satélite. Si un planeta es ligeramente desplazado de su órbita, dándosele alguna pequeña velocidad, se plantea esta cuestión: ¿oscilará alrededor de su órbita para volver tal vez a ella después de cierto tiempo, o se alejará de ella? En el primer caso la órbita es estable, y en el último inestable. Así, la pregunta de si el movimiento primario de los planetas o cualesquiera irregularidades en sus movimientos son periódicos es vital.

Como sabemos (cap. 21, sec. 7), Lagrange había encontrado soluciones periódicas especiales en el problema de los tres cuerpos. No

se llegó a nuevas soluciones periódicas de dicho problema hasta que George William Hill (1838-1914), el primer gran matemático americano, realizó sus trabajos sobre la teoría lunar. En 1877, Hill publicó en privado un ensayo notablemente original sobre el movimiento del perigeo de la Luna [47]. También publicó un artículo muy importante sobre el movimiento de la Luna en el *American Journal of Mathematics* [48]. Su trabajo fundó la teoría matemática de las ecuaciones diferenciales lineales homogéneas con coeficientes periódicos.

La primera idea importante de Hill (en su ensayo de 1877) consistió en determinar una solución periódica de las ecuaciones diferenciales para el movimiento de la Luna, que se aproximaba al movimiento factual observado. Después formuló ecuaciones para las variaciones a partir de esta solución periódica, lo que condujo a un sistema de cuarto orden de ecuaciones diferenciales lineales con coeficientes periódicos. Conociendo algunas integrales, fue capaz de reducir su sistema de cuarto orden a una única ecuación diferencial lineal de segundo orden

$$\frac{d^2x}{dt^2} + \theta(t)x = 0, \tag{26}$$

con $\theta(t)$ periódica de período π y par. La forma de la ecuación de Hill puede ser escrita, expandiendo $\theta(t)$ en serie de Fourier, como

$$\frac{d^2x}{dt^2} + x(q_0 + 2q_1 \cos 2t + 2q_2 \cos 4t + \ldots) = 0. \tag{27}$$

Hill llamó $\zeta = e^{it}$, $q_{-\alpha} = q_\alpha$ y escribió (27) como

$$\frac{d^2x}{dt^2} + x \sum_{-\infty}^{\infty} q_\alpha \zeta^{2\alpha} = 0. \tag{28}$$

Más adelante escribió

$$x = \sum_{j=-\infty}^{\infty} b_j \zeta^{\mu+2j},$$

[47] Esto se reimprimió en *Acta Math.*, 8, 1886, 1-36 = *Coll. Math. Works*, 1, 243-270.

[48] *Amer. Jour. of Math.*, 1, 1878, 5-26, 129-147, 245-260 = *Coll. Math. Works*, 1, 284-335.

donde μ y b_j tenían que ser determinados. Sustituyendo este valor de x en (28) y haciendo los coeficientes de cada potencia de ζ iguales a 0, obtuvo un sistema de ecuaciones lineales doblemente infinito

$$\begin{array}{r}
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\ldots[2]b_{-2} - q_1b_{-1} - q_2b_0 - q_3b_1 - q_4b_2 - \ldots = 0 \\
\ldots - q_1b_{-2} + [-1]b_{-1} - q_1b_0 - q_2b_1 - q_3b_2 - \ldots = 0 \\
\ldots - q_2b_{-2} - q_1b_{-1} + [0]b_0 - q_1b_1 - q_2b_2 - \ldots = 0 \\
\ldots - q_bb_{-2} - q_2b_{-1} - q_1b_0 + [1]b_1 - q_1b_2 - \ldots = 0 \\
\ldots - q_4b_{-2} - q_3b_{-1} - q_2b_0 - q_1b_1 + [2]b_2 - \ldots = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}$$

donde

$$[j] = (\mu + 2j)^2 - q_0.$$

Hill hizo el determinante de los coeficientes de las incógnitas b_j igual a 0. Primero determinó las propiedades de las infinitas soluciones para μ y proporcionó fórmulas explícitas para determinar las μ. Con estos valores de μ, resolvió el sistema de una infinidad de ecuaciones homogéneas lineales en las infinitas b_j para la razón de las b_j a b_0. Hill demostró que la ecuación diferencial de segundo orden tiene una solución periódica y que el movimiento del perigeo de la Luna es periódico.

El trabajo de Hill fue menospreciado hasta que Poincaré [49] demostró la convergencia del procedimiento y, de este modo, colocó a la teoría de los determinantes infinitos y sistemas infinitos de ecuaciones lineales sobre una base sólida. La atención de Poincaré y los esfuerzos finales de Hill dieron importancia a éste, ayudando también a la materia en cuestión.

8. Ecuaciones diferenciales no lineales: la teoría cualitativa

Poincaré inició, bajo el estímulo del trabajo de Hill, un nuevo enfoque en la búsqueda de soluciones periódicas para ecuaciones

---

[49] *Bull. Soc. Math. de France*, 13, 1885, 19-27; 14, 1886, 77-90 = *Œuvres*, 5, 85-94, 95-107.

diferenciales gobernando los movimientos planetarios y la estabilidad de las órbitas planetarias y de los satélites. Puesto que las ecuaciones relevantes son no lineales, Poincaré estudió esta clase. Las ecuaciones diferenciales ordinarias no lineales prácticamente habían aparecido desde el principio de la materia como, por ejemplo, la ecuación de Riccati (cap. 21, sec. 4), la ecuación del péndulo y las ecuaciones de Euler del cálculo de variaciones (cap. 24, sec. 2), aunque no se desarrollaron métodos generales para su solución.

En vista del hecho de que las ecuaciones del movimiento de incluso tres cuerpos no pueden ser resueltas explícitamente en términos de funciones conocidas, el problema de estabilidad no puede ser resuelto examinando la solución. Por lo tanto, Poincaré buscó métodos mediante los cuales el problema sería resuelto examinando las propias ecuaciones diferenciales. La teoría iniciada por él se llamó teoría cualitativa de las ecuaciones diferenciales. La presentó en cuatro artículos, todos bajo el mismo título: «Mémoire sur les courbes définies par una équation différentielle» («Memoria sobre las curvas definidas por una ecuación diferencial»)[50]. Las preguntas que buscaba contestar fueron enunciadas por él mismo con estas palabras: «¿Describe una curva cerrada el punto que se mueve? ¿Se mantiene siempre en el interior de cierta porción del plano? En otras palabras, y hablando en el lenguaje de la astronomía, nos hemos preguntado si la órbita es estable o inestable.»

Poincaré empezó con ecuaciones no lineales de la forma

$$\frac{dy}{dx} = \frac{P(x, y)}{Q(x, y)}, \tag{29}$$

donde P y Q son analíticas en x e y. Esta forma fue escogida en parte porque algunos problemas del movimiento planetario lo condujeron hasta ahí, y porque era el sistema matemático más sencillo con el cual comenzar el tipo de investigacion que Poincaré tenía a la vista. La solución de (29) es de la forma $f(x,y) = 0$, y se dice que esta ecuación define un sistema de trayectorias. En lugar de $f(x,y) = 0$ uno puede considerar la forma paramétrica $x = x(t)$, $y = y(t)$.

En el análisis de las clases de soluciones que puede tener la ecuación (29), Poincaré encontró que los puntos singulares de la ecuación

[50] *Jour. de Math.* (3), 7, 1881, 375-422; 8, 1882, 251-296; (4), 1, 1885, 167-244; 2, 1886, 151-217 = *Œuvres*, 1, 3-84, 90-161, 167-221.

diferencial (los puntos en que ambas P y Q se anulan) juegan un papel importante. Estos puntos singulares son indeterminados o irregulares en el sentido de Fuchs. Aquí Poincaré usó trabajos previos de Briot y Bouquet (sec. 5), pero se limitó a valores reales y a estudiar el comportamiento de la solución entera en lugar de circunscribirse a un entorno de los puntos singulares. Distinguió cuatro tipos de puntos singulares y describió el comportamiento de las soluciones alrededor de tales puntos.

El primer tipo de punto singular es el foco *(foyer)*, el origen en la Figura 29.2, y la solución describe una espiral alrededor y se aproxima al origen cuando va de $-\infty$ a ∞. Este tipo de solución es considerada estable. El segundo tipo de punto singular es el punto de ensilladura *(col.)*. Es el origen de la Figura 29.3 y las trayectorias se aproximan a este punto y luego se apartan de él. Las líneas AA' y BB' son las asíntotas de las trayectorias. El movimiento es inestable. El tercer tipo de punto singular, llamado nodo *(nœud)*, es un punto donde se cruzan una infinidad de soluciones, y el cuarto, llamado centro, es uno a cuyo alrededor existen trayectorias cerradas, unas encerrando otras y todas encerrando al centro.

Entre otros muchos resultados, Poincaré encontró que pueden existir curvas cerradas que no *tocan* a ninguna de las curvas que satisfacen la ecuación diferencial. Llamó a estas curvas cerradas ciclos sin contacto. Una curva que satisface la ecuación diferencial no

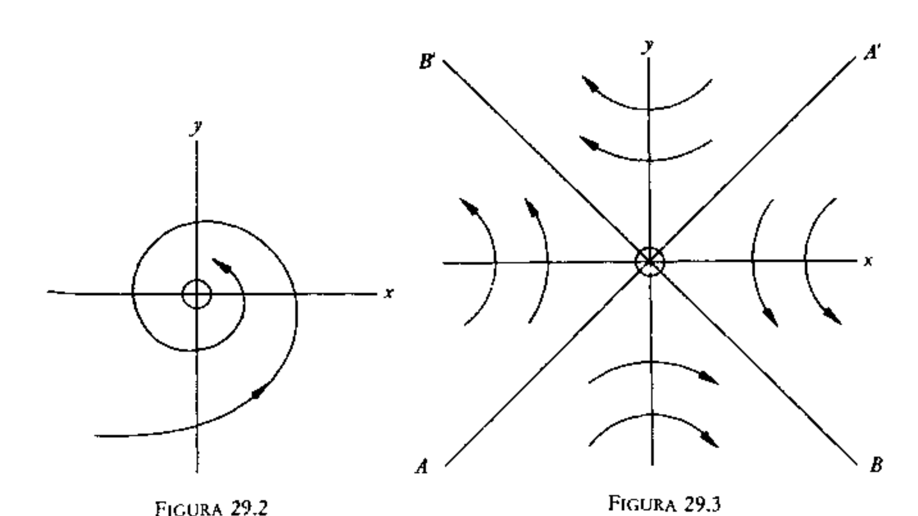


FIGURA 29.2

FIGURA 29.3

puede tocar tal ciclo en más de un punto, y si cruza el ciclo no lo puede volver a cruzar. Tal curva, si es la órbita de un planeta, representa un movimiento inestable.

Además de los ciclos sin contacto, existen curvas cerradas que Poincaré llamó ciclos límite. Estas son curvas cerradas que satisfacen la ecuación diferencial y a las que otras soluciones se aproximan asintóticamente, esto es, sin llegar nunca al ciclo límite. El acercamiento puede ser a partir de fuera o de dentro del ciclo límite C (Fig. 29.4). Para algunas ecuaciones diferenciales del tipo (29) determinó los ciclos límite y las regiones donde existen. En el caso de los ciclos límite, las trayectorias aproximan una curva periódica y el movimiento es de nuevo estable. Sin embargo, si la dirección de los movimientos fuera la de apartarse del ciclo límite, el movimiento exterior sería inestable y el movimiento interior una espiral en contracción.

En el tercero de los artículos sobre esta materia, Poincaré estudió ecuaciones de primer orden de grado superior y de la forma $F(x,y,y') = 0$, donde F es un polinomio en x, y e y'. Para estudiar estas ecuaciones, Poincaré consideró x, y e y' como tres coordenadas cartesianas y consideró la superficie definida por la ecuación diferencial. Si esta superficie tiene género 0 (forma de una esfera), entonces las curvas integrales tienen las mismas propiedades que en el caso de las ecuaciones diferenciales de primer grado. Para otros géneros, los resultados sobre las curvas integrales pueden ser bastante diferentes. Así, para un toro surgen muchas nuevas circunstancias. Poincaré no completó su estudio. En el cuarto artículo (1886), estudió las ecuaciones de segundo orden y obtuvo algunos resultados análogos a los de las ecuaciones de primer orden.

Mientras continuaba sus trabajos sobre los tipos de soluciones de la ecuación diferencial (29), Poincaré consideró una teoría más general dirigida al problema astronómico de los tres cuerpos. En un ensayo premiado, «Sur le problème de trois corps et les équations de la dynamique» («Sobre el problema de los tres cuerpos y las ecuaciones de la dinámica»)[51], consideró el sistema de ecuaciones diferenciales

$$\frac{dx_i}{dt} = X_i(x_1, \ldots, x_n, \mu), \qquad i = 1, 2, \ldots, n. \qquad (30)$$

[51] *Acta Math.*, 13, 1890, 1-270 = *Œuvres*, y, 262-479.

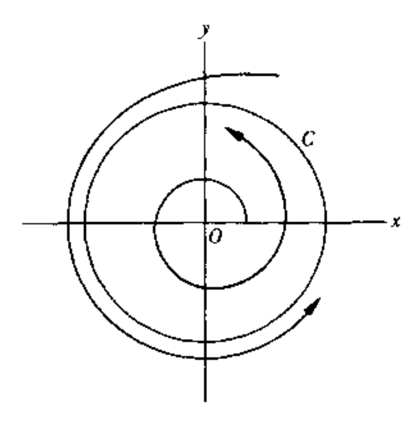


FIGURA 29.4

Desarrolló las X_i en potencias del parámetro pequeño μ y, suponiendo que el sistema tenía para $\mu = 0$ una solución periódica conocida

$$x_i = \phi_i(t), \qquad i = 1, 2, ..., n$$

con período T, propuso encontrar la solución periódica del sistema que para $\mu = 0$ se reduce a $\phi_i(t)$. La existencia de soluciones periódicas para el problema de los tres cuerpos ya la había descubierto Hill, y Poincaré hizo uso de este hecho.

Los detalles del trabajo de Poincaré son demasiado especializados para considerarlos aquí. Primero generalizó trabajos anteriores de Cauchy sobre soluciones de sistemas de ecuaciones diferenciales ordinarias donde este último utilizó su *calcul des limites*. Luego, Poincaré demostró la existencia de las soluciones periódicas que buscaba y aplicó lo aprendido en el estudio de las soluciones periódicas del problema de los tres cuerpos para el caso donde las masas de dos de los cuerpos (pero no la del Sol) son pequeñas. De esta forma se obtienen tales soluciones suponiendo que las dos pequeñas masas se mueven en círculos concéntricos alrededor del Sol y están en el mismo plano. Se pueden obtener otras suponiendo que para $\mu = 0$, las órbitas son elipses y que sus períodos son conmensurables. Con estas soluciones, y usando la teoría que había desarrollado para este sistema, obtuvo otras soluciones periódicas. En suma, demostró que existe una infinidad de posiciones iniciales y velocidades iniciales tales que las *distancias mutuas* de los tres cuerpos son funciones

periódicas del tiempo. (Tales soluciones también son llamadas periódicas.)

En su artículo de 1890, Poincaré dedujo muchas otras conclusiones acerca de las soluciones periódicas y casi periódicas del sistema (30). Entre ellas está el muy notable descubrimiento para tal sistema de una nueva clase de soluciones previamente desconocida. Las denominó soluciones asintóticas. Existen dos clases. En la primera, la solución aproxima la solución periódica asintóticamente cuando t aproxima $-\infty$ o cuando t aproxima $+\infty$. La segunda clase consiste en soluciones doblemente asintóticas: soluciones que aproximan una solución periódica cuando t aproxima a $-\infty$ y $+\infty$. Hay una infinidad de tales soluciones doblemente asintóticas. Todos los resultados de este artículo de 1890 y muchos otros están en *Les Méthodes nouvelles de la mécanique céleste (Los nuevos métodos de la mecánica celeste)* de Poincaré [52].

El trabajo de Poincaré sobre la estabilidad del sistema solar tuvo éxito sólo parcialmente. La estabilidad es aún un problema abierto. De hecho, también lo es la cuestión respecto a si la órbita de la Luna es estable; muchos científicos piensan que no lo es.

La estabilidad de las soluciones de (29) se analiza mediante la ecuación característica, a saber,

$$\begin{vmatrix} Q_x(x_0, y_0) - \lambda & P_x(x_0, y_0) \\ Q_y(x_0, y_0) & P_y(x_0, y_0) - \lambda \end{vmatrix} = 0, \tag{31}$$

donde (x_0,y_0) es un punto singular de (29). La estabilidad en un entorno de (x_0,y_0), de acuerdo con el teorema del distinguido matemático ruso Alexander Liapunoff (1857-1918), depende de las raíces de esta ecuación característica [53]. El análisis de los casos posibles es detallado e incluye muchos más tipos de los que se vieron en la discusión del trabajo de Poincaré. El resultado básico, de acuerdo con Liapunoff, cuyo análisis en torno a problemas de estabilidad continuó hasta los primeros años de este siglo, es que las soluciones son estables en la vecindad de un punto singular cuando y sólo cuando las raíces de la ecuación (31) en λ tienen partes reales negativas.

[52] Tres volúmenes, 1892-1899.

[53] *Ann. Fac. Sci. de Toulouse* (2), 9, 1907, 203-474; publicado originalmente en ruso en 1892.

El estudio cualitativo de las ecuaciones no lineales avanzó mediante la introducción por Poincaré de argumentos topológicos (en el primero de los cuatro artículos en el *Journal de Mathématiques*). Para describir la naturaleza de un punto singular introdujo la noción de índice. Considérese un punto singular P_0 y la curva cerrada simple C rodeándolo. En cada intersección de C con las soluciones de

$$\frac{dy}{dx} = \frac{P(x, y)}{Q(x, y)} \tag{32}$$

existe un ángulo de dirección de trayectoria, que denotaremos por ϕ, el cual puede tener cualquier valor de 0 a 2 π radianes. Si un punto se mueve ahora en direccion contraria a las manecillas del reloj alrededor de C (Fig. 29.5), el ángulo ϕ variará; y después de completar el circuito alrededor de C, ϕ tendrá el valor $2\pi I$ donde I es un entero o cero (ya que el ángulo de dirección de las trayectorias ha regresado a su valor inicial). La cantidad I es el índice de la curva. Puede ser demostrado que el índice de una curva cerrada que contiene varias singularidades es la suma algebraica de sus índices. El índice de una trayectoria cerrada es $+1$ e inversamente.

La naturaleza de las trayectorias se determina por la ecuación característica, y de forma que el índice I de una curva puede ser determinado sólo por el conocimiento de la ecuación diferencial. Se puede demostrar que

$$I = \frac{1}{2\pi} \int_C d\left(\arc \tan \frac{P}{Q}\right) = \frac{1}{2\pi} \int_C \frac{P\,dQ - Q\,dP}{P^2 + Q^2},$$

donde la trayectoria de integración es la curva cerrada C.

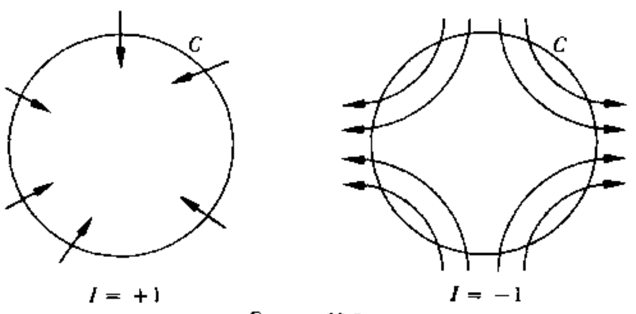


FIGURA 29.5

Después de Poincaré, el trabajo más significativo sobre soluciones de ecuaciones de la forma (32) se debe a Ivar Bendixson (1861-1935). Uno de sus resultados [54] principales da un criterio mediante el cual, en ciertas regiones, se demuestra que no existe ninguna trayectoria cerrada. Sea D una región donde $\partial Q/\partial x + \partial P/\partial y$ tiene el mismo signo. Entonces, la ecuación (32) no tiene solución periódica en D.

El teorema ahora llamado en honor de Poincaré y Bendixson, que está contenido en el artículo último de 1901, proporciona un criterio positivo para la existencia de una solución periódica de (32). Si P y Q están definidas y son regulares en $-\infty < x$, $y < \infty$ y si cuando t se aproxima a ∞ una solución $x(t)$, $y(t)$ permanece dentro de una región acotada del plano (x,y) sin aproximarse a puntos singulares, entonces existe al menos una curva solución cerrada de la ecuación diferencial.

El estudio de ecuaciones no lineales que inició Poincaré fue ensanchado en varias direcciones. Se mencionará también otro tema iniciado en el siglo XIX. Las ecuaciones diferenciales lineales que estudió Fuchs poseen la propiedad que los puntos singulares están fijos y, de hecho, son determinados por los coeficientes de las ecuaciones diferenciales. En el caso de ecuaciones no lineales los puntos singulares pueden variar con las condiciones iniciales y son llamados puntos singulares movibles (o móviles). Así, la ecuación $y' + y^2 = 0$ tiene la solución general $y = 1/(x - c)$ donde c es arbitraria. La localización de la singularidad en la solución depende del valor de c. El fenómeno de los puntos singulares movibles lo descubrió Fuchs [55]. El estudio de los puntos singulares movibles y las ecuaciones de segundo orden no lineales con y sin tales puntos singulares fue abordado por muchos científicos, en especial, por Paul Painlevé (1863-1933). Una característica interesante es que muchos de los tipos de ecuaciones de segundo orden de la forma $y'' = f(x,y,y')$ requieren para su solución nuevos tipos de funciones trascendentes, ahora llamadas trascendentes de Painlevé [56].

El interés por las ecuaciones no lineales se ha incrementado en el siglo XX. Las aplicaciones han ido de la astronomía a problemas de comunicaciones, servomecánica, sistemas de control automático

---

[54] *Acta Math.*, 24, 1901, 1-88.

[55] *Sitzungsber. Akad. Wiss. zu Berlin*, 1884, 699-710 = *Werke*, 2, 364 sigs.

[56] *Comp. Rend.*, 143, 1906, 1111-1117.

y electrónico. El estudio también se ha desplazado del aspecto cualitativo a las investigaciones cuantitativas.

Bibliografía

Acta Mathematica, vol. 38, 1921. Este volumen está dedicado en su totalidad a artículos sobre la obra de Poincaré por varios grandes matemáticos.

Bocher, M.: «Randwertaufgaben bei gewohnlichen Differentialgleichungen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, II, A7a, 437-463.

— : «Boundary problems in one dimension», *Internat. Cong. of Math., Proc.,* Cambridge, 1912, 1, 163-195.

Burkhardt, H.: Entwicklungen nach oscillirenden Functionen und Integration der Differential gleichungen der mathematischen Physik». *Jahres. der Deut. Math.-Verein*, 10, 1908, 1-1804.

Craig, T.: «Some of the developments in the Theory of Ordinary Differential Equations between 1878 and 1893». *N. Y. Math. Soc. Bull.*, 2, 1893, 119-134.

Fuchs, Lazarus: *Gesammelte mathematische Werke*, 3 vols., 1904-1909, Georg Olms (reimpresión), 1970.

Heine, Eduard: *Handbuch der Kugelfunktionen,* 2 vols., 1878-1881, Physika Verlag (reimpresión) 1961.

Hilb, E.: «Lineare Differentialgleichungen im komplexen Gebiet», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, II, B5.

— : «Nichtlineare Differentialgleichungen», *Encyk. der Math. Wiss.,* B. G. Teubner, 1899-1916, II, B6.

Hill, George W.: *Collected Mathematical Works*, 4 vols., 1905, Johnson Reprint Corp., 1965.

Klein, Felix: *Vorlesungen uber Entwicklung der Mathematik im 19. Jahrhundert*, vol. I, Chelsea (reimpresión), 1950.

— : *Gesammelte mathematische Abhandlungen*, Julius Springer, 1923, vol. 3.

Painlevé, P.: «Le Problème moderne de l'intégration des équations différentielles», *Third Internat. Math. Cong. in Heidelberg,* 1904, 86-99, B. G. Teubner, 1905.

— : «Gewohnliche Differentialgleichungen, Existenz der Losungen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, II, A4a.

Poincaré, Henri: *Œuvres*, 1, 2 y 5, Gauthier-Villars, 1928, 1916 y 1960.

Riemann, Bernhard: *Gesammelte mathematische Werke*, 2.ª ed., 1892, Dover (reimpresión), 1953.

Schlesinger, L.: «Bericht uber die Entwickelung der Theorie der linearen Differentialgleichungen seit 1865», *Jahres. der Deut. Math.-Verein.*, 18, 1909, 133-266.

Wangerin, A.: «Theorie der Kugelfunktionen und der verwandten Funktionen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1899-1916, II, A10.
Wirtinger, W.: «Riemanns Vorlesungen uber die hypergeometrische Reihe und ihre Bedeutung», *Third Internat. Math. Cong. in Heidelberg*, 1904, B. G. Teubner, 1905, 121-139.

Capítulo 30
EL CALCULO DE VARIACIONES EN EL SIGLO XIX

Aunque penetrar en los misterios íntimos de la naturaleza y aprender desde allí las causas verdaderas de los fenómenos no nos sea permitido, puede suceder, sin embargo, que ciertas hipótesis ficticias pueden ser suficientes para explicar muchos fenómenos.

LEONHARD EULER

1. Introducción

Como hemos visto, el cálculo de variaciones fue fundado esencialmente por Euler y Lagrange durante el siglo XVIII. Más allá de los problemas físico-matemáticos de diversas clases, existía una motivación fundamental para su estudio, a saber, el Principio de Mínima Acción, que en manos de Maupertuis, Euler y Lagrange se convirtió en el principio básico de la física matemática. Los científicos del siglo XIX continuaron su labor sobre la acción mínima y el gran estímulo del cálculo de variaciones, a lo largo de la primera mitad del siglo, vino de esta dirección. Físicamente, el interés radicaba en la ciencia de la mecánica y particularmente en problemas de astronomía.

2. La física matemática y el cálculo de variaciones

La exitosa formulación de Lagrange hizo de las leyes de la dinámica en términos de su Principio de Acción Mínima, sugirió que la

idea debería ser aplicada a otras ramas de la física. Lagrange [1] proporcionó un principio de mínimo para la dinámica de fluidos (aplicable a fluidos compresibles y no compresibles) a partir del cual derivó las ecuaciones de Euler para la mecánica de fluidos (cap. 22, sec. 8) y se jactó con claridad al decir que un principio de mínimo gobernaba este campo como lo hacía con el movimiento de partículas y de cuerpos rígidos. También se resolvieron muchos problemas de elasticidad con el cálculo de variaciones en la primera parte del siglo XIX por Poisson, Sophie Germain, Cauchy y otros; también este trabajo ayudó a mantener el campo activo, aunque no se vieron nuevas ideas matemáticas del cálculo de variaciones en esta área o en la famosa contribución de Gauss a la mecánica: el Principio de Mínima Restricción [2].

La primera novedad digna de mención respecto a este punto se debe a Poisson, que valiéndose de las coordenadas generalizadas de Lagrange, siguió inmediatamente sus dos ensayos para empezar con [3] las ecuaciones de Lagrange (cap. 24, sec. 5)

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} = 0, \qquad i = 1, 2, \ldots, n. \tag{1}$$

Aquí la energía kinética T expresada en coordenadas generalizadas es $2T = \sum_{i,j=1}^{n} a_{ij}\dot{q}_i\dot{q}_j$, V es la energía potencial y T y V son independientes de t. Hace $L = T - V$, donde V depende solamente de las q_i y no de las $\dot{q}_i$, por lo que puede escribir

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i}, \tag{2}$$

de tal forma que las ecuaciones del movimiento son

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \ldots, n. \tag{3}$$

También introduce

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i}, \tag{4}$$

[1] *Misc. Taur.*, 2_2, 1760-1761, 196-298, pub. 1762 = *Œuvres*, 1, 365-468.

[2] *Jour. für Math.*, 4, 1829, 232-235 = *Werke*, 5, 23-28.

y a partir de (3) tiene

$$\dot{p}_i = \frac{\partial L}{\partial q_i}, \quad i = 1, 2, \ldots, n. \tag{5}$$

Las p_i son las componentes del momento, mientras que las q_i son coordenadas rectangulares de posición. La ecuación (5) es un paso en la dirección que debemos ver ahora.

A William R. Hamilton se debe el gran cambio en la formulación de los principios de acción mínima, importante para el cálculo de variaciones y las ecuaciones diferenciales parciales. Hamilton llegó a la dinámica por medio de la óptica. Su meta en física consistió en formar una estructura matemática deductiva a la manera del tratamiento de la mecánica de Lagrange.

Hamilton también empezó con un principio de acción mínima y se proponía deducir otros nuevos. Sin embargo, su actitud hacia tales principios difería profundamente de la de Maupertuis, Euler y Lagrange. En un ensayo publicado en el *Dublin University Review* [4] dice: «Pero, aunque la ley de acción mínima ha obtenido así un rango entre los más altos teoremas de la física, sin embargo sus pretensiones de una necesidad cosmológica, sobre la base de la economía en el universo, por lo común son rechazados hoy. Y justo el rechazo aparece, entre otras razones, porque la cantidad que se pretendía economizar con frecuencia es gastada suntuosamente.» Ya que en algunos fenómenos de la naturaleza, incluso muy sencillos, la acción es maximizada, Hamilton prefería hablar de un principio de acción estacionaria.

En una serie de artículos del período de 1824 a 1832, Hamilton desarrolló su teoría matemática de la óptica y transfirió ideas que había introducido allí a la mecánica. Escribió dos ensayos básicos [5]. El más pertinente es el segundo de ellos. Aquí introduce la integral de acción, a saber, la integral en el tiempo de la diferencia entre las energías cinética y potencial

$$S = \int_{P_1,t_1}^{P_2,t_2} (T - V)\, dt. \tag{6}$$

[3] *Jour. de l'Ecole Poly.*, 8, 1809, 266-344.

[4] 1833, 795-826 = *Math. Papers*, 1, 311-332.

[5] *Phil. Trans.*, 1834, Part II, 247-308; 1835, Part I, 95-144 = *Math. Papers*, 2, 103-211.

La cantidad $T - V$ es llamada la función lagrangiana, aunque no obstante la introdujera Poisson; P_1 representa a $q_1^1, q_2^1, ..., q_n^1$ y P_2 a $q_1^{(2)}, q_2^{(2)}, ..., q_n^{(2)}$. Luego, Hamilton generaliza el principio de Euler y Lagrange al permitir trayectorias de comparación que no están restringidas, excepto en que el movimiento a lo largo de ellas deba empezar en P_1 en el tiempo t_1 y finalizar en P_2 al tiempo t_2. También se tiene que la ley de la conservación de la energía no necesita cumplirse, mientras que en el principio de Euler-Lagrange la conservación de la energía está presupuesta y, como consecuencia, el tiempo requerido por un objeto para recorrer cualquiera de las trayectorias de comparación difiere del tiempo tomado para recorrer la trayectoria efectiva.

El principio hamiltoniano de acción mínima asevera que el movimiento es de hecho el que hace la acción estacionaria. Para sistemas conservativos, esto es, donde las componentes de fuerza son derivables de un potencial que es función únicamente de la posición, $T + V =$ const. De aquí que, $T - V = 2T -$ const., y el principio de Hamilton se reduce al de Lagrange, pero, como se notó, el principio de Hamilton también se mantiene para sistemas no conservativos. Asimismo, la energía potencial V puede ser una función del tiempo y aun de las velocidades; esto es, en coordenadas generalizadas $V = V(q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n, t)$.

Si, al escribir $T - V$ igual a L, escribimos la integral de acción (6) como

$$S = \int_{t_1}^{t_2} L(q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n, t)\, dt, \qquad (7)$$

con la condición de que todas las comparaciones $q_i(t)$ deben tener los mismos valores dados en t_1 y t_2, entonces el problema es determinar las q_i como funciones de t a partir de la condición que las verdaderas q_i hacen la integral estacionaria. Las ecuaciones de Euler, que expresan la condición que la primera variación de S es 0, se convierten en un sistema de ecuaciones diferenciales ordinarias de segundo orden simultáneas, a saber,

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = 0, \quad k = 1, 2, ..., n, \qquad (8)$$

y las ecuaciones deben ser resueltas en $t_1 \leqslant t \leqslant t_2$. Estas ecuaciones

aún hoy son llamadas ecuaciones lagrangianas de movimiento, a pesar de que L es ahora una función diferente. La elección del sistema de coordenadas es arbitrario y utiliza comúnmente coordenadas generalizadas. Esta es una ventaja esencial de los principios variacionales.

Ahora introduce (véase (4))

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

Entonces las ecuaciones (8) se convierten en

$$\dot{p}_i = \frac{\partial L}{\partial q_i}.$$

La introducción de las p_i como un nuevo conjunto de variables independientes suele ser atribuido a Hamilton, aunque fue hecha primero por Poisson. Tenemos ahora el sistema simétrico de ecuaciones diferenciales

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad \dot{p}_i = \frac{\partial L}{\partial q_i}, \quad i = 1, 2, ..., n. \tag{9}$$

Este es un sistema de $2n$ ecuaciones diferenciales ordinarias de primer orden en p_i y $\dot{p}_i$. Sin embargo, las $\dot{p}_i$ son dp_i/dt.

En su segundo artículo (1835), Hamilton simplifica este sistema de ecuaciones. Introduce una nueva función H, la cual está definida mediante

$$H(p_i, q_i, t) = -L + \sum_{i=1}^{n} p_i\dot{q}_i. \tag{10}$$

Esta función es físicamente la energía total, ya que se puede mostrar que la sumatoria es igual a $2T$. La transformación de L a H es llamada transformación de Legendre, ya que fue usada por Legendre en su trabajo sobre ecuaciones diferenciales ordinarias. El que H sea una función de las p_i, q_i y t, mientras que $L = T - T$ sea una función de las q_i, $\dot{q}_i$ y t, resulta del hecho de que, dado que $p_i = \partial L/\partial \dot{q}_i$ podemos resolver para las $\dot{q}_i$ y sustituir en L.

Con (10), se puede mostrar que las ecuaciones diferenciales del movimiento (9) toman la forma

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad i = 1, 2, ..., n. \tag{11}$$

Se supone que la función H es conocida en la aplicación a problemas físicos. Estas ecuaciones son un sistema de $2n$ ecuaciones diferenciales ordinarias de primer orden en las $2n$ variables dependientes p_i y q_i, como funciones de t, mientras que las ecuaciones de Lagrange (1) son un sistema de n ecuaciones de segundo orden en las $q_i(t)$. Más tarde, Jacobi llamó a las ecuaciones de Hamilton las ecuaciones diferenciales canónicas. Son las ecuaciones variacionales (ecuaciones de Euler) para la integral

$$S = \int_{P_1,t_1}^{P_2,t_2} (T - V)\, dt = \int L\, dt = \int \left\{ \sum_{i=1}^{n} p_i \dot{q}_i - H(p_i, q_i, t) \right\} dt.$$

Este conjunto de ecuaciones aparece en uno de los artículos de Lagrange de 1809 que trata de la teoría de perturbaciones para sistemas mecánicos. Sin embargo, mientras que Lagrange no reconoció la conexión básica de estas ecuaciones con las ecuaciones del movimiento, Cauchy sí lo hizo en un artículo inédito de 1831. En 1835, Hamilton hizo de estas ecuaciones la base de sus investigaciones sobre mecánica.

Para usar las ecuaciones de Hamilton del movimiento es posible frecuentemente expresar H en un sistema apropiado de coordenadas p y q de tal forma que el sistema de ecuaciones (11) sea resoluble para las p_i y q_i como funciones del tiempo. En particular, si podemos seleccionar coordenadas de tal forma que H dependa únicamente de las p_i, el sistema es resoluble.

En un artículo de 1837 [6], y en conferencias sobre dinámica de los años de 1842 y 1843, que fueron publicadas en 1866 en el clásico *Vorlesungen über Dynamik* (Lecciones sobre dinámica), Jacobi demostró que el proceso de Hamilton es susceptible de invertirse. En la teoría de Hamilton, si se conoce la acción S o el hamiltoniano H, es posible formar las $2n$ ecuaciones diferenciales canónicas e intentar

[6] *Jour. für Math.*, 17, 1837, 97-162 = *Ges. Werke*, 4, 57-127.

dar solución al sistema. La idea de Jacobi era encontrar coordenadas P_j y Q_j de tal forma que H sea tan simple como sea posible, y entonces las ecuaciones diferenciales (11) serían integradas fácilmente. Específicamente, buscó la transformación

$$\begin{aligned} Q_j &= Q_j(p_i, q_i, t) \\ P_j &= P_j(p_i, q_i, t) \end{aligned} \tag{12}$$

de tal forma que

$$\delta \int_{t_1}^{t_2} \left(\sum_{i=1}^{n} p_i \dot{q}_i - H(p_i, q_i, t) \right) dt = 0$$

se convierte mediante la transformación (12) en

$$\delta \int_{t_1}^{t_2} \left(\sum_{i=1}^{n} P_i \dot{Q}_i - K(P_i, Q_i, t) \right) dt = 0,$$

y así las ecuaciones diferenciales del hamiltoniano son

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}, \tag{13}$$

donde $K(P_i, Q_i, t)$ es el nuevo hamiltoniano. Esta trayectoria lleva a

$$K = H(p_i, q_i, t) + \frac{\partial \Omega}{\partial t}(Q_i, q_i, t),$$

donde Ω es una nueva función, llamada función generadora de la transformación. Jacobi eligió $K = 0$ de tal forma que por (13)

$$\dot{Q}_i = 0, \qquad \dot{P}_i = 0,$$

y Q_i y P_i son constantes. Más aún,

$$H + \frac{\partial \Omega}{\partial t} = 0, \tag{14}$$

y puede demostrarse que

$$p_i = \frac{\partial \Omega}{\partial q_i}.$$

De aquí, por (14), en vista de las variables en H,

$$H\left(\frac{\partial\Omega}{\partial q_i}, q_i, t\right) + \frac{\partial\Omega}{\partial t}(Q_i, q_i, t) = 0. \qquad (15)$$

Ya que $\dot{Q}_i = 0$, $Q_i = \alpha_i$, y así la ecuación es de primer orden en Ω con las variables independientes q_i y t. Con este cambio la ecuación (15) es la ecuación diferencial parcial de Hamilton-Jacobi para la función Ω. Si esta ecuación puede ser resuelta para una Ω completa, esto es, una conteniendo n constantes arbitrarias, la solución tendría la forma

$$\Omega(\alpha_1, \alpha_2, ..., \alpha_n, q_1, q_2, ..., q_n, t).$$

Ahora bien, es un hecho de la teoría de transformaciones de Jacobi que

$$P_i = -\frac{\partial\Omega}{\partial\alpha_i}$$

y que $P_i = \beta_i$, constante, porque $\dot{P}_i = 0$. Entonces se busca la solucion de las ecuaciones algebraicas

$$\frac{\partial\Omega}{\partial\alpha_i} = -\beta_i, \quad i = 1, 2, ..., n$$

para las q_i. Estas soluciones

$$q_i = f_i(\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_n, t), \quad i = 1, 2, ..., n,$$

son las soluciones de las ecuaciones canónicas de Hamilton. De este modo Jacobi había mostrado que se puede resolver el sistema de ecuaciones ordinarias (11) resolviendo la ecuación diferencial parcial (15). El mismo Jacobi encontró la Ω necesaria para muchos problemas de la mecánica.

La aportación de Hamilton representó la culminación de una serie de esfuerzos encaminados a hallar un principio amplio a partir del cual podían ser derivadas las leyes del movimiento de varios problemas de la mecánica. Inspiró la lucha por obtener principios variacionales similares en otras ramas de la física matemática, tales como la elasticidad, teoría electromagnética, relatividad y teoría cuán-

tica. Los principios que han sido derivados, aun el principio de Hamilton, no son necesariamente las aproximaciones más prácticas a la solución de problemas particulares. Más bien, el atractivo de tales formulaciones generales se apoya en intereses filosóficos y estéticos, aunque los hombres de ciencia ya no infieren que la existencia de un principio máximo-mínimo sea una prueba de la sabiduría y eficiencia de Dios.

Desde el punto de vista de la historia de las matemáticas, el trabajo de Hamilton y Jacobi es significativo porque motivó más investigación no solamente en el cálculo de variaciones, sino también sobre sistemas de ecuaciones diferenciales ordinarias y ecuaciones diferenciales parciales de primer orden.

3. Extensiones matemáticas del propio cálculo de variaciones

Los resultados de Euler y Legendre ofrecían nada más condiciones necesarias (cap. 24), pues, por ejemplo, en el caso más sencillo, para maximizar o minimizar la integral

$$J = \int_a^b f(x, y, y')\, dx, \qquad (16)$$

durante cerca de cincuenta años, después de la obra de Legendre, los matemáticos hicieron nuevas exploraciones de las primeras y segundas variaciones, pero aun así no se obtuvieron resultados decisivos. En 1837 [7], Jacobi encontró cómo hacer más precisa la condición de Legendre, de tal forma que condujera a una condición suficiente. Su principal descubrimiento en este sentido fue el concepto de punto conjugado. Veamos primero a qué se refiere esto.

Considérense las curvas que satisfacen la ecuación (característica) de Euler; tales curvas son llamadas extremales. Para el problema básico del cálculo de variaciones, existe una familia de extremales pasando por un punto dado *A*. Supóngase ahora que *A* es uno de los extremos entre los que buscamos una curva maximante o minimizante. Dada cualquier otra extremal, el punto límite de la intersección de otras extremales conforme se acercan al extremal es el punto conjugado de *A* sobre ese extremal. Otra manera de verlo es

[7] *Jour. für Math.*, 17, 1837, 68-82 = *Ges. Werke*, 4, 39-55.

decir que tenemos una familia de curvas y éstas posiblemente tienen a su vez una envolvente. El punto de contacto de cualquier extremal y la envolvente de la familia es el punto conjugado de A sobre ese extremal. Entonces la condición de Jacobi es que si $y(x)$ es un extremal entre los puntos extremos A y B del problema original, ningún punto conjugado debe estar sobre la extremal entre A y B ni ser el mismo B.

Lo que esto significa de una manera concreta se verá a través de un ejemplo. Puede mostrarse que la trayectoria parabólica es, de todas las trayectorias (Fig. 30.1) saliendo de A con velocidad constante v, pero con ángulos variantes de fuego, la extremal del problema de maximizar o minimizar la integral de acción

$$\frac{m}{2}\int_A^B v\, ds.$$

El problema de minimizar la acción entre dos puntos A y B tiene en general dos soluciones, la parábola $AA''B$ y la parábola ABA'. Sucede también que la familia de parábolas que pasa por A tiene una envolvente que toca las dos parábolas en A'' y A'. El punto conjugado sobre $AA''B$ es A'' y sobre ABA' es A'. De acuerdo con la condición de Jacobi, el extremal $AA''B$ no puede suministrar un máximo o un mínimo; sin embargo, el extremal ABA' sí puede.

Jacobi reconsideró la segunda variación $\delta^2 J$ (cap. 24, sec. 4). Si escribimos $y + \epsilon t(x)$ en lugar de la $y + \delta y$ de Lagrange y si a y b son las abscisas de A y B entonces

$$\delta^2 J = \frac{\varepsilon^2}{2}\int_a^b (t^2 f_{yy} + 2tt' f_{yy'} + t'^2 f_{y'y'})\, dx. \tag{17}$$

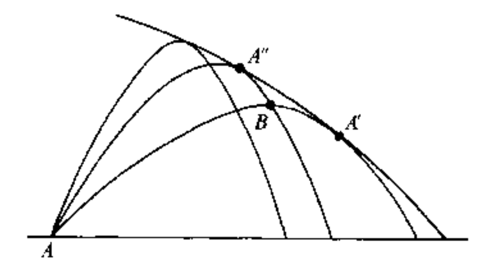


FIGURA 30.1

Jacobi demostró que

$$\delta^2 J = \frac{\varepsilon^2}{2} \int_a^b f_{y'y'} \left(t' - t \frac{u'}{u} \right)^2 dx,$$

donde u es una solución de la ecuacion accesoria de Jacobi

$$\left\{ f_{yy} - \frac{d}{dx} f_{yy'} \right\} u - \frac{d}{dx} (f_{y'y'} u') = 0 \tag{18}$$

y donde las derivadas parciales son calculadas a lo largo del extremal uniendo los dos puntos extremos A y B. Ahora se requiere que $u(x)$ pase por A. Entonces todos los otros puntos sobre el extremal $y(x)$ a través de A y B en los cuales se anula $u(x)$ son los puntos conjugados de A sobre ese extremal. Si $u = \beta_1 u_1 + \beta_2 u_2$ es la solución general de la ecuación accesoria (18), entonces se demuestra que

$$\frac{u_1(x)}{u_2(x)} = \frac{u_1(a)}{u_2(a)},$$

donde a es la abscisa del punto A, es la ecuación de las abscisas de los puntos conjugados de A.

Jacobi demostró también que no se necesita resolver la ecuación accesoria. Dado que se debe resolver la ecuación de Euler en cualquier caso, sea $y = y(x, c_1, c_2)$ la solución general de esa ecuación, esto es, la familia de extremales. Entonces se tomaría u_1 como $\partial y / \partial c_1$, y u_2 como $\partial y / \partial c_2$.

Jacobi dedujo dos conclusiones a partir de este trabajo sobre puntos conjugados. La primera fue que si a lo largo del extremal de A a B hay un punto conjugado de A, entonces es imposible un máximo o un mínimo. Jacobi estaba esencialmente en lo correcto con esta conclusión.

Sobre la base de sus consideraciones sobre puntos conjugados, Jacobi concluyó además que un extremal (una solución de la ecuación de Euler) tomado entre A y B para la cual $f_{y'y'} > 0$ a lo largo de la curva y que no tiene ningún punto conjugado entre A y B (o en B) suministra un mínimo para la integral original. El argumento correspondiente con $f_{y'y'} < 0$ —según aseveró— se mantiene para un máximo. De hecho, estas condiciones suficientes no eran correctas, como veremos dentro de unos momentos. En este ensayo de 1837,

Jacobi enunció resultados y proporcionó indicaciones breves de las demostraciones. Las demostraciones completas de los enunciados correctos fueron suministradas por otros estudiosos de épocas posteriores.

Aparte del valor específico de los resultados de Jacobi para la existencia de una función maximizadora o minimizadora, su trabajo hizo evidente que el progreso en el cálculo de variaciones no podía ser guiado por la teoría de máximos y mínimos del cálculo ordinario.

Durante treinta y cinco años ambas conclusiones de Jacobi se aceptaron como correctas. Durante este período los artículos sobre la materia eran imprecisos respecto a los enunciados y dudosos en las demostraciones; había poca claridad en la formulación de los problemas y contenían errores de todo tipo. Entonces, Weierstrass realizó su trabajo sobre el cálculo de variaciones. Presentó sus resultados durante las conferencias que dio en Berlín en 1872, pero no las publicó él mismo. Sus ideas motivaron un nuevo interés, impulsando mayor actividad en la materia y afinando el pensamiento, tal como lo hiciera el trabajo de Weierstrass en otros dominios.

El primer punto importante de Weierstrass era que los criterios hasta entonces establecidos para un mínimo o un máximo —los de Euler, Legendre y Jacobi— eran limitados porque la supuesta curva minimizadora o maximizadora $y(x)$ se comparaba con otras curvas $y(x) + et(x)$, donde de hecho se daba por supuesto que $et(x)$ y $et'(x)$, o lo que Lagrange llamó δy y $\delta y'$, eran pequeñas a lo largo del rango de x de A a B. Esto es, $y(x)$ estaba siendo comparada con una clase limitada de otras curvas, y en cuanto a la satisfacción de los tres criterios lo hacía mejor que cualquier otra de *estas* curvas de comparación. A tales variaciones se las denominó variaciones débiles por Adolf Kneser (1862-1930). Sin embargo, para encontrar la curva que realmente maximiza o minimiza la integral J, uno debe compararla con *todas* las otras curvas uniendo A con B, para incluir aquellas cuyas derivadas puedan no aproximar las derivadas de la curva maximizante (o minimizante), del modo como las curvas de comparación se acercan en posición a la curva maximizante. De esta manera, una curva de comparación puede tener una esquina aguda (Fig. 30.2) en uno o varios lugares a lo largo del rango de x de A a B. Las curvas de comparación consideradas por Weierstrass son las que Kneser llamó variaciones fuertes.

Weierstrass demostró en 1879 que para las variaciones débiles las tres condiciones —que la curva sea un extremal (una solución de la

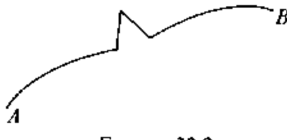


FIGURA 30.2

ecuación de Euler), que $f_{y'y'} > 0$ a lo largo del extremal y que cualquier punto conjugado de A debe caer más allá de B— son claramente condiciones suficientes para que el extremal produzca un mínimo de la integral J ($f_{y'y'} < 0$ para un máximo).

Más adelante, Weierstrass consideró variaciones fuertes. Para estas variaciones, introdujo primero una cuarta condición necesaria: definió una nueva función llamada función E, o función exceso, mediante

$$E(x, y, y', \bar{p}) = f(x, y, \bar{p}) - f(x, y, y') - (\bar{p} - y')f_{y'}(x, y, y'), \qquad (19)$$

y su resultado fue: la cuarta condicion necesaria para que $y(x)$ proporcione un mínimo es que $E(x,y,y',\bar{p}) \geqslant 0$ a lo largo del extremal $y(x)$ para todo valor finito de $\bar{p}$. Para un máximo $E \leqslant 0$.

Entonces (en 1879), Weierstrass dirigió su atención a las condiciones suficientes para un máximo (o un mínimo) cuando se permiten variaciones fuertes. Para formular sus condiciones suficientes es necesario introducir el concepto de campo de Weierstrass. Considérese cualquier familia de un parámetro (Fig. 30.3) de extremales $y = \Phi(x,\gamma)$ en la que el extremal particular uniendo A y B está incluido, digamos para $\gamma = \gamma_0$. Independientemente de algunos detalles sobre la continuidad y diferenciabilidad de $\Phi(x,\gamma)$, el hecho esencial acerca de esta familia de extremales es que en una región alre-

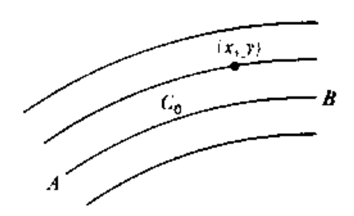


FIGURA 30.3

dedor del extremal entre A y B, pasa a través de cualquier punto (x,y) de la región uno (y solamente uno) extremal de la familia. Una familia de extremales que cumple esta condición es llamada un campo.

Dado un campo rodeando el extremal C_0 que une A y B (Fig. 30.4), entonces, si en cada punto (x,y) entre $x = a$ y $x = b$ y en la región cubierta por el campo, $E(x,y,p(x,y),\tilde{p}) \geqslant 0$ donde $p(x,y)$ denota la pendiente en (x,y) del extremal pasando por (x,y) y $\tilde{p}$ es cualquier valor finito, entonces C_0 minimiza la integral J con respecto a cualquier otra C dentro del campo y uniendo A y B. (Para un máximo, $E \leqslant 0$).

En 1900, Hilbert[8] introdujo su teoría de la integral invariante, la cual simplificó grandemente la condición de suficiencia. Hilbert hizo la pregunta: ¿es posible determinar la función $p(x,y)$ de tal forma que la integral

$$I = \int_{x_1}^{x_2} \left\{ f(x, y, p) - (y' - p) f_{y'}(x, y, p) \right\} dx \qquad (20)$$

sea independiente de la trayectoria en una región de valores de (x,y)? Hilbert encontró que si $p(x,y)$ está determinada de este modo, entonces las soluciones de la ecuación diferencial

$$\frac{dy}{dx} = p(x, y)$$

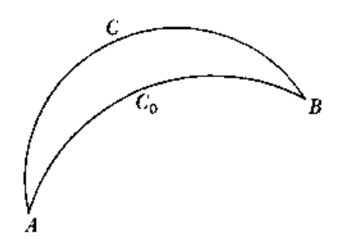


FIGURA 30.4

[8] *Nachrischten König. Ges. der Wiss. zu Gött.*, 1900, 291-296 = *Ges. Abh.*, 3, 323-329. Existe una traducción al inglés por Mary Winston Newson en el *Amer. Math. Soc. Bull.*, 8, 1902, 472-478. Este material es parte del famoso artículo de 1900, «Problemas matemáticos».

son extremales de un campo. Inversamente, si $p(x,y)$ es la función pendiente de un campo F, entonces I es independiente de la trayectoria en F. A partir de este teorema Hilbert derivó la condición de suficiencia de Weierstrass para variaciones fuertes.

4. Problemas relacionados en el cálculo de variaciones

Nuestra exposición de la historia del cálculo de variaciones ha estado concentrada intensamente en la integral

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx.$$

Se ha hecho una mención de otros problemas: los problemas isoperimétricos, los problemas de varias funciones de una variable tales como los que surgen en el Principio de Acción Mínima y el caso de integrales múltiples —que Lagrange trató por primera vez y aparecen en el problema de las superficies mínimas (cap. 24, sec. 4)—. Hay muy numerosos tipos de problemas relacionados, tales como aquellos en los que la curva minimizadora o maximizadora es tratada con una representación paramétrica $x = x(t)$ e $y(t)$ —este problema fue ampliamente discutido por Weierstrass— y problemas que fundamentalmente se dan en la dinámica, donde las variables que aparecen en el integrando están restringidas por ecuaciones auxiliares o subsidiarias, llamadas restricciones. El último tipo de problema está de algún modo relacionado con el problema isoperimétrico, ya que también existe una condición subsidiaria, a saber, la longitud de la curva encerrando el área máxima está fijada aunque en ese problema la condición subsidiaria se encuentra en la forma de una integral que expresa la longitud de la curva, mientras que en el caso de las restricciones dinámicas la condición subsidiaria (o condiciones) están en la forma de ecuaciones donde entran las variables independientes o dependientes e incluso las diferenciales de las variables dependientes. Además subyace el problema básico denominado problema de Dirichlet, discutido anteriormente (cap. 28, secs. 4 y 8).

No rastrearemos la historia detallada de estos problemas porque no surgió de ahí ninguna característica básica de las matemáticas, aunque los problemas son significativos y se haya realizado considerable trabajo sobre ellos hasta el presente. Es tal vez digno de

mención el que el problema de las superficies mínimas, que requiere resolver la ecuación

$$(1 + q^2)r - 2pqs + (1 + p^2)t = 0$$

había estado dormitando después del artículo de Ampère de 1817 hasta que el físico belga Jóseph Plateau (1801-1883) en un libro de 1873, *Statique expérimentale et théorique des liquides soumis aux seules formes moléculaires (Estática experimental y teórica de los líquidos sometido únicamente a sus formas moleculares)*, mostró que si uno sumerge cables teniendo la forma de curvas cerradas dentro de una solución de glicerina (o agua jabonosa) y entonces los retira, una película de jabón que tiene la forma de la superficie de área mínima, se apoyará en el borde del alambre. Así los matemáticos recibieron nuevos impulsos para considerar superficies mínimas acotadas por una curva cerrada en el espacio. Ya que la curva o curvas frontera pueden ser bastante complicadas, de hecho la solución explícita analítica para la superficie mínima sería imposible de obtener. Este problema, ahora conocido como el problema de Plateau, condujo a trabajar sobre la demostración de al menos la existencia de las soluciones, de las que pueden ser deducidas algunas propiedades de las soluciones.

Bibliografía

Dresden, Arnold: «Some recent work in the Calculus of Variations», *Amer. Math. Soc. Bull.*, 32, 1926, 475-521.
Duren, W. L., Jr.: «The development of sufficient conditions in the Calculus of Variations», *University of Chicago Contributions to the Calculus of Variations,* I, 1930, 245-349, University of Chicago Press, 1931.
Hamilton, W. R.: *The Mathematical Papers*, 3 vols., Cambridge University Press, 1931, 1940 y 1967.
Jacobi, C. G. J.: *Gesammelte Werke,* G. Reimer, 1886 y 1891, Chelsea (reimpresión), 1968, vols. 4 y 7.
— : *Vorlesungen über Dynamik* (1866), Chelsea (reimp.), 1968. También en vol. 8 de los *Gesammelte Werke* de Jacobi.
McShane, E. J.: «Recent developments in the Calculus of Variations», *Amer. Math. Soc. Semicentennial Publications,* II, 1938, 69-97.
Porter, Thomas Isaac: «A history of the classical isoperimetric problem»,

University of Chicago Contributions to the Calculus of Variations, II, 475-517, University of Chicago Press, 1933.

Prange, Georg: «Die allgemeinen Integrationsmethoden der analytischen Mechanik», *Encyk. der Math. Wiss.*, B. G. Teubner, 1904-1935, IV, 2, 509-804.

Todhunter, Isaac: *A History of the Calculus of Variations in the Nineteenth Century*, Chelsea (reimpresión), 1962.

Weierstrass, Karl: *Werke*, Akademische Verlagsgesellschaft, 1927, vol. 7.

Capítulo 31
LA TEORIA DE GALOIS

> Desafortunadamente, lo que se reconoce poco es que los libros científicos más valiosos son aquellos en los cuales el autor indica claramente lo que él no sabe, porque en general un autor impresiona a sus lectores escondiendo las dificultades.
>
> EVARISTE GALOIS

1. Introducción

El centro escénico del álgebra a principios del siglo XIX lo ocupaba la solución de ecuaciones polinominales. Durante este período, Galois resuelve definitiva y comprensiblemente el problema general referente a qué ecuaciones se resolvían mediante operaciones algebraicas. No solamente creó el primer cuerpo coherente de teoría algebraica, sino que además introdujo nuevas nociones que habrían de desarrollarse en otras teorías algebraicas con mayor posibilidad de aplicación. En particular, los conceptos de grupo y cuerpo surgieron de sus investigaciones y las del Abel.

2. Ecuaciones binómicas

Ya hemos discutido (cap. 25, sec. 2) los esfuerzos inútiles de Euler, Vandermonde, Lagrange y Ruffini para resolver algebraicamente ecuaciones de grado mayor que 4 y la ecuación binomial $x^n - 1 = 0$. Gauss obtuvo un gran éxito. En la última sección de sus *Disquisi-*

tiones Arithmeticae [1], Gauss consideró la ecuación

$$x^p - 1 = 0, \tag{1}$$

donde p es primo [2]. A menudo dicha ecuación es llamada ecuación ciclotómica o ecuación de la división del círculo. El último término se refiere al hecho de que sus raíces son, por el teorema de De Moivre,

$$x_j = \cos\frac{k2\pi\theta}{p} + i \operatorname{sen}\frac{k2\pi\theta}{p}, \qquad k = 1, 2, ..., p, \tag{2}$$

y los números complejos x_j, cuando se dibujan geométricamente son los vértices de un polígono regular de p lados que se encuentran sobre el círculo unitario.

Gauss demostró que las raíces de esta ecuación pueden expresarse racionalmente en términos de las raíces de una serie de ecuaciones

$$Z_1 = 0, Z_2 = 0, ..., \tag{3}$$

cuyos coeficientes son racionales en las raíces de las ecuaciones precedentes de la sucesión. Los grados de la ecuación (3) son precisamente los factores primos de $p - 1$. Existe un Z_i para cada factor, incluso si está repetido. Además, cada una de los $Z_i = 0$ se resolverá por radicales y así la ecuación (1) también puede resolverse por radicales.

Este resultado es, por supuesto, de importancia primordial para el problema de resolver algebraicamente la ecuación general de grado n. Demuestra que algunas ecuaciones de grado alto se resuelven por radicales; por ejemplo, una ecuación de quinto grado si 5 es un factor de $p - 1$ o una ecuación de grado séptimo si 7 es tal factor.

El resultado es también de gran importancia para el problema geométrico de construir polígonos regulares de p lados. Si $p - 1$ no contiene otros factores más que 2, entonces el polígono es construible con regla y compás, ya que los grados de la ecuación (3) son siempre 2 y cada una de sus raíces es construible en términos de sus

---

[1] 1801, *Werke*, 1.

[2] El caso de p primo incluye $x^n - 1 = 0$, ya que si $n = pq$, sea $y = x^q$. Pero $y^p - 1 = 0$ es soluble. De aquí que $x^q =$ const. puede ser resuelta si q es primo y si no q puede ser descompuesto de la misma manera que n.

coeficientes. Así, estamos en condiciones de construir todos los polígonos de un número primo de lados p si $p - 1$ es una potencia de 2. Tales primos son 3, 5, 17, 257, 65537, ... Alternativamente, es posible que un polígono regular sea construido si p es un primo de la forma $2^{2^h} + 1$[3]. Gauss indica (art. 365) que a pesar de que la construcción geométrica de polígonos regulares de 3, 5 y 15 lados, y de aquellos que se derivan inmediatamente de ellos —por ejemplo, 2^n, $2^n.3$, $2^n.5$, $2^n.15$, donde n es un entero positivo— era conocida en el tiempo de Euclides, en un intervalo de 2000 años no se habían descubierto nuevos polígonos construibles y los geómetras fueron unánimes al declarar que no se podía construir otros.

Gauss pensó que su resultado llevaría a todo tipo de intentos por encontrar nuevos polígonos construibles con un número primo de lados. Entonces advierte: «Siempre que $p - 1$ contiene otros factores primos además de 2, llegamos a ecuaciones superiores, a saber, a una o más ecuaciones cúbicas si 3 entra una vez o con más frecuencia como factor de $p - 1$, a ecuaciones del quinto grado si $p - 1$ es divisible por cinco, etc. Y podemos demostrar con todo rigor que estas ecuaciones superiores no pueden ser evitadas o hechas depender de ecuaciones de grado inferior; y aunque los límites de este trabajo no nos permiten dar una demostración aquí, seguimos pensando que es necesario indicar este hecho y decir que no se debe seguir buscando construir otros polígonos [de un número primo de lados] que aquellos proporcionados por nuestra teoría, como por ejemplo, polígonos de 7, 11, 13 y 19 lados, y así emplear este tiempo en vano.»

Más adelante, Gauss considera (art. 366) polígonos de cualquier número de lados n y asevera que un polígono regular de n lados es construible si y sólo si $n = 2^l\, p_1.p_2,..., p_n$, donde los $p_1, p_2,...,p_n$ son primos distintos de la forma $2^{2^h} + 1$ y donde l es cualquier entero positivo o 0. La suficiencia de esta condición se sigue fácilmente del trabajo de Gauss sobre polígonos de un número primo de lados pero la necesidad no es tan obvia y no fue demostrada por Gauss[4].

La construcción de polígonos regulares había interesado a Gauss

[3] En un primo de la forma $2^{\mu} + 1$, μ es necesariamente de la forma 2^h, pero $2^{2^h} + 1$ no es primo necesariamente.

[4] Véase: James Pierpont, «On an undemostrated theorem of the Disquisitiones Arithmeticae», *Amer. Math. Soc. Bull.*, 2, 1895-1896, 77-83. Este artículo proporciona la demostración. El hecho de que la condición de Gauss es necesaria fue demostrado primero por Pierre L. Wantzel (1814-1848), *Jour. de Math.*, 2, 1837, 366-372.

desde 1796, cuando concibió su primera demostración de que un polígono de 17 lados es construible. Existe una historia acerca de este descubrimiento y es útil su repetición. Este problema de construcción ya era famoso: un día Gauss se acercó a su profesor A. G. Kästner, de la Universidad de Göttingen, con la demostración de que este polígono es construible. Kästner no creía en tal aseveración y buscó deshacerse de Gauss, como hoy en día los profesores despiden a los trisectores de ángulos. En lugar de tomar su tiempo para examinar la demostración de Gauss y encontrar el supuesto error en ella, Kästner le dijo a Gauss que la construcción carecía de importancia, ya que todas las construcciones prácticas eran conocidas. Por supuesto, Kästner sabía que la existencia de construcciones prácticas o aproximadas era irrelevante para el problema teórico. Para interesar a Kästner en su demostración, Gauss le señaló que había resuelto una ecuación algebraica de grado diecisiete. Kästner respondió que la solución era imposible. Pero Gauss respondió que había reducido el problema a resolver una ecuación de menor grado. «Oh, está bien», se burló Kästner, «yo ya he hecho lo mismo». Más tarde Gauss le pagó a Kästner, quien también se enorgullecía de su poesía, elogiando a Kästner como el mejor poeta entre los matemáticos y el mejor matemático entre los poetas.

3. El trabajo de Abel sobre la solución de ecuaciones por radicales

Abel leyó el trabajo de Lagrange y Gauss sobre la teoría de ecuaciones y cuando aún estudiaba el grado de preparatoria atacó el problema de la solubilidad de ecuaciones de grado superior siguiendo el tratamiento de Gauss de la ecuación binómica. Primero, Abel pensó que había resuelto la ecuación general de quinto grado por radicales. Pero pronto, convencido de su error, intentó evidenciar que tal demostración no era posible (1824-1826). Primero tuvo éxito en demostrar el teorema: las raíces de una ecuación soluble por radicales pueden ser dadas de tal forma que cada uno de los radicales que aparece en las expresiones para las raíces es expresable como una función racional de las raíces de la ecuación y ciertas raíces de la unidad. Abel usó entonces su teorema para demostrar [5] la impo-

[5] *Jour. für Math.*, 1, 1826, 65-84 = *Œuvres*, 1, 66-94.

sibilidad de resolver por radicales la ecuación general de grado mayor que cuatro.

La demostración de Abel, realizada ignorando el trabajo de Ruffini (cap. 25, sec. 2), es indirecta e innecesariamente complicada. Su artículo también contenía un error en una clasificación de funciones, que afortunadamente no era esencial para el argumento. Más tarde publicó otras dos demostraciones más elaboradas. Una demostración simple, sencilla y rigurosa basada en las ideas de Abel la proporcionó Kronecker en 1879 [6].

Así, el problema de la solución de ecuaciones generales de grado mayor que cuatro fue zanjada por Abel. También consideró algunas ecuaciones especiales. Tomó [7] el problema de la división de la lemniscata (resolver $x^n - 1 = 0$ es el equivalente al problema de la división del círculo en n arcos iguales) y llegó a una clase de ecuaciones algebraicas, ahora llamadas ecuaciones abelianas, resolubles por radicales. La ecuación ciclotómica (1) es un ejemplo de ecuación abeliana. Más generalmente, una ecuación es abeliana si todas sus raíces son funciones racionales de una de ellas, esto es, si las raíces son x_1, $\theta_1(x_1)$, $\theta_2(x_1),\ldots,\theta_{n-1}(x_1)$ donde las θ_i son funciones racionales. Existe también la condición que $\theta_\alpha(\theta_\beta(x_1)) = \theta_\beta(\theta_\alpha(x_1))$ para todos los valores de α y β de 1 a $n-1$.

En este último trabajo introdujo dos conceptos (aunque no la terminología): cuerpo y polinomio irreducible en un cuerpo dado. Por cuerpo de números, como Gauss más tarde, dio a entender una colección de números tales que la suma, diferencia, producto y cociente de dos números cualesquiera en la colección (excepto la división por 0) están también en la colección. Así, los números racionales, los números reales y los números complejos forman un cuerpo. Se dice que un polinomio es reducible en un cuerpo (comúnmente el cuerpo al cual pertenecen los coeficientes) si puede ser expresado como el producto de dos polinomios de grados menores y con coeficientes en el cuerpo. Si el polinomio no puede ser expresado de tal forma se dice que es irreducible.

Abel atacó entonces el problema de caracterizar todas las ecuaciones que son resolubles por radicales y había comunicado algunos

---

[6] *Monatsber. Berliner Akad.*, 1879, 205-229 = *Werke*, 4, 73-96. La demostración de Kronecker es explicada por James Pierpont en «On the Ruffini-Abelian Theorem», *Amer. Math. Soc. Bull.*, 2, 1895-1896, 200-221.

[7] *Jour. für Math.*, 4, 1829, 131-156 = *Œuvres*, 1, 478-507.

resultados a Crelle y a Legendre poco antes de que la muerte se lo llevara en 1829.

4. La teoría de resolubilidad de Galois

Después de la obra de Abel la situación era como sigue: a pesar de que la ecuación general de grado superior a cuatro no era soluble por radicales, había muchas ecuaciones especiales, tales como las ecuaciones binómicas $x^p = a$, p primo, y las ecuaciones abelianas que eran solubles por radicales. La finalidad ahora era determinar qué ecuaciones son solubles por radicales. Esta tarea justamente iniciada con Abel, la adoptó Evariste Galois (1811-1832). Nacido de padres holgados y bien educados, asistió a uno de los renombrados liceos de París e inició su estudio de las matemáticas a los quince años de edad. Esta materia se convirtió en su pasión y se abocó con empeño a estudiar los trabajos de Lagrange, Gauss, Cauchy y Abel. Ignoró las otras materias. Galois quiso entrar en la Ecole Polytechnique, pero posiblemente porque falló en explicar con suficiente detalle las preguntas que tenía que contestar oralmente en el examen de admisión, o tal vez porque los examinadores no lo entendieron, fue rechazado en dos intentos. Por tales circunstancias, entró a la Ecole Préparatoire (era el nombre entonces de la Ecole Normale, y era escuela muy inferior por aquellos días). Durante la revolución de 1830, que arrojó del trono a Charles X e instaló a Louis Philippe, Galois criticó públicamente al director de la escuela por haber dejado de apoyar la revolución y este hecho produjo su expulsión. Fue arrestado en dos ocasiones por ofensas políticas, y pasó casi el último año y medio de su vida en prisión, de donde sale para morir en un duelo el 31 de mayo de 1832.

Durante su primer año en la Escuela, Galois publicó cuatro artículos. En 1829 presentó dos artículos sobre la solución de ecuaciones a la Academia de Ciencias. Estos le fueron confiados a Cauchy, quien los perdió. En enero de 1830 presentó a la Academia otro escrito cuidadosamente redactado sobre sus investigaciones, el cual se le envió a Fourier, quien murió poco después, y dicho ensayo también se perdió. Siguiendo la sugerencia de Poisson, Galois (1831) escribió un nuevo ensayo sobre sus investigaciones. Este texto, «Sur les conditions de résolubilité des équations par radicaux» («Sobre las

condiciones de resolubilidad de las ecuaciones por radicales») [8], su único artículo completo sobre su teoría de la solución de ecuaciones, le fue devuelto por Poisson como ininteligible, con la recomendación de que debería escribir una explicación más amplia. La noche anterior a su duelo, Galois bosquejó precipitadamente un apresurado resumen de sus investigaciones que confió a un amigo, August Chevalier. Este resumen se ha conservado.

En 1846, Liouville editó y publicó en el *Journal de Mathématiques* [9] parte de los artículos de Galois, incluyendo una revisión de su artículo de 1831. Más adelante, el *Cours d'algèbre supérieure* (*Curso de álgebra superior*, 3.ª ed.) de Serret de 1866, proporcionó una exposición de las ideas de Galois. Camile Jordan, en 1870, en su libro *Traité des substitutions et des équations algébraiques (Tratado de las sustituciones y de las ecuaciones algebraicas)*, hace la primera presentación clara y completa de la teoría de Galois.

Galois estudió el problema de caracterizar las ecuaciones solubles por radicales mejorando las ideas de Lagrange, aunque también derivó algunas sugerencias del trabajo de Legendre, Gauss y Abel. Propuso considerar la ecuación general, la cual es, por supuesto

$$x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0 \quad (4)$$

donde, como en el trabajo de Lagrange, los coeficientes deben ser independientes o arbitrarios por completo, y ecuaciones particulares, tales como

$$x^4 + px^2 + q = 0, \quad (5)$$

donde únicamente dos coeficientes son independientes. El pensamiento central de Galois consistió en evitar la construccion de los resolventes de Lagrange (cap. 25, sec. 2) de la ecuación polinómica dada, una construcción que requiere gran habilidad y que no tiene una metodología clara.

Como Lagrange, Galois hace uso de la notación de sustituciones o permutaciones de las raíces. Así, si x_1, x_2, x_3 y x_4 son las cuatro raíces de una ecuación de cuarto grado, el intercambio de x_1 y x_2 en cualquier expresión en las x_i es una sustitución. Esta sustitución

[8] *Œuvres*, 1897, 33-50.
[9] *Jour. de Math.*, 11, 1846, 381-444.

particular está indicada por

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_3 & x_4 \end{pmatrix}.$$

Una segunda sustitución está indicada mediante

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_1 & x_2 \end{pmatrix}.$$

Realizar la primera sustitución y más adelante la segunda es equivalente a efectuar la tercera sustitución

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_3 & x_1 & x_2 \end{pmatrix}.$$

porque, por ejemplo, en la primera sustitución x_1 es reemplazada por x_2; y mediante la segunda sustitución x_2 es reemplazada por x_4; y por la tercera sustitución x_1 va directamente a x_4. Uno dice que el *producto* de las dos primeras sustituciones tomadas en el orden así indicado es la tercera sustitución. Existen en total 4! sustituciones posibles. Se dice que el conjunto de sustituciones forma un grupo porque el producto de dos sustituciones cualesquiera es un miembro del conjunto. Esta noción, que por supuesto no es una definición formal de grupo abstracto, se debe a Galois.

Para asegurar algún entendimiento de las ideas de Galois consideremos la ecuación [10]

$$x^4 + px^2 + q = 0,$$

donde p y q son independientes. Sea R el cuerpo formado por las expresiones racionales en p y q con coeficientes en el cuerpo de los números racionales, una expresión típica es $(3p^2 - 4q) / (q^2 - 7p)$. Se dice con Galois que R es el cuerpo obtenido al añadir las letras o indeterminadas p y q a los números racionales. Este cuerpo R es el cuerpo o dominio de racionalidad de los coeficientes de la ecua-

[10] Ya que la propia presentación de las ideas de Galois no fue clara e introdujo tantas nociones nuevas, utilizaremos un ejemplo de Verriest (véase la bibliografía al final del capítulo) para aclarar la teoría de Galois.

ción dada y se dice que la ecuación pertenece al cuerpo R. Como Abel, Galois no usó los términos cuerpo o dominio de racionalidad, pero sí usó este concepto.

Sucede que nosotros sabemos que las raíces de la ecuación de cuarto grado son

$$x_1 = \sqrt{\frac{-p + \sqrt{p^2 - 4q}}{2}}, \qquad x_2 = -\sqrt{\frac{-p + \sqrt{p^2 - 4q}}{2}}$$

$$x_3 = \sqrt{\frac{-p - \sqrt{p^2 - 4q}}{2}}, \qquad x_4 = -\sqrt{\frac{-p - \sqrt{p^2 - 4q}}{2}}.$$

Entonces es cierto que las dos relaciones con coeficientes en R

$$x_1 + x_2 = 0 \qquad \text{y} \qquad x_3 + x_4 = 0$$

se cumplen para las raíces. Ya que nuestra ecuación dada es de cuarto grado, hay veinticuatro posibles sustituciones de las raíces. Las ocho substituciones siguientes

$$E = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix}, \qquad E_1 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_3 & x_4 \end{pmatrix},$$

$$E_2 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & x_4 & x_3 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_4 & x_3 \end{pmatrix},$$

$$E_4 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_1 & x_2 \end{pmatrix}, \qquad E_5 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_3 & x_1 & x_2 \end{pmatrix},$$

$$E_6 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_2 & x_1 \end{pmatrix}, \qquad E_7 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_3 & x_2 & x_1 \end{pmatrix},$$

dan las dos relaciones verdaderas en R [11]. Es conveniente saber que

[11] Para la ecuación general de grado n, esto es, con n cantidades independientes como coeficientes, una función de las raíces es *invariante* bajo (o inalterada por) la substitución de las raíces si y sólo si permanece idéntica a la función original. Si los coeficientes son todos numéricos, entonces la función es inalterada si permanece la misma numéricamente. Así, para $x^3 + x^2 + x + 1 = 0$ las raíces son $x_1 = -1$, $x_2 = i$ y $x = -i$. Considérese x_2^2. La sustitución de x_3 por x_2 da x_3^2. Esta tiene el mismo

estas ocho son las únicas sustituciones de las veinticuatro que dejan invariantes *todas* las relaciones en R entre las raíces, son el grupo de la ecuación en R, un subgrupo del grupo total. Esto es, el grupo de una ecuación con respecto a un cuerpo R es el grupo o subgrupo de sustituciones de las raíces que dejan invariantes *todas* las relaciones con los coeficientes en R entre las raíces de una ecuación dada (ya sea general o particular). Se puede decir que el número de sustituciones que dejan todas las relaciones en R invariantes es una medida de nuestra ignorancia de las raíces, ya que no podemos distinguirlas bajo estas ocho sustituciones.

Ahora considérese $x_1^2 - x_3^2$, que es igual a $\sqrt{p^2 - 4q}$. Adjuntamos este radical a R, y formamos el cuerpo R', esto es, formamos el cuerpo más pequeño conteniendo R y $\sqrt{p^2 - 4}\, q$. Entonces

$$x_1^2 - x_3^2 = \sqrt{p^2 - 4q} \qquad (6)$$

es una relación en R'. Ya que $x_1 + x_2 = 0$ y $x_3 + x_4 = 0$ también tenemos que

$$x_1^2 = x_2^2 \qquad \text{y} \qquad x_3^2 = x_4^2.$$

Entonces, a la vista de los dos últimos hechos podemos decir que las cuatro primeras de las ocho sustituciones anteriores hacen la relación (6) en R' verdadera, pero las últimas cuatro no. Entonces las cuatro sustituciones, si dejan invariante *toda* relación verdadera en R' entre las raíces, son el grupo de la ecuación en R'. Las cuatro son un subgrupo de las ocho.

Ahora supongamos que adjuntamos a R' la cantidad $\sqrt{(-p-D)/2}$, donde $D = \sqrt{p^2 - 4q}$ y se forma el cuerpo R''. Entonces

$$x_3 - x_4 = 2\sqrt{\frac{-p - D}{2}}$$

es una relación en R''. Esta relación permanece invariante sólo por las primeras dos sustituciones E y E_1 pero no por el resto de las

valor numérico que x_2^2. Entonces x_2^2 no es alterada por la sustitución. Si los coeficientes contienen algunos valores numéricos y algunas cantidades independientes entonces una función de las raíces permanece inalterada por una sustitución de las raíces si la función permanece numéricamente la misma para todos los valores de las cantidades independientes (en el dominio en el que éstas puedan estar limitadas) y los valores numéricos que las raíces puedan tomar.

ocho. Entonces el grupo de la ecuación en R'' consiste de estas dos sustituciones, teniendo en cuenta que cada relación en R'' entre las raíces permanece invariante bajo estas dos sustituciones. Las dos son un subgrupo de las cuatro sustituciones.

Si ahora adjuntamos a R'' la cantidad $\sqrt{(-p + D)/2}$ obtenemos R'''. En R''' tenemos

$$x_1 - x_2 = 2\sqrt{\frac{-p + D}{2}}\,.$$

Ahora la única sustitución que deja todas las relaciones en R''' verdaderas es E y este es el grupo de la ecuación en R'''.

Podemos ver a partir de la discusión anterior que el grupo de una ecuación es clave para su solubilidad, ya que el grupo expresa el grado de indistinción de las raíces. Nos dice lo que desconocemos acerca de las raíces.

Había muchos grupos, o más estrictamente un grupo de sustituciones y subgrupos sucesivos, relacionados con lo anterior. Ahora el *orden* del grupo (o subgrupo) es el número de elementos contenidos en él. De esta forma, tenemos grupos de órdenes 24, 8, 4, 2 y 1. El orden de un subgrupo divide siempre el orden del grupo (sec. 6). El *índice* de un subgrupo es el orden del grupo en el cual está, dividido por el orden del subgrupo. Así, el índice de un subgrupo de orden ocho es tres.

El esbozo anterior muestra meramente las ideas con las que trataba Galois. Su trabajo procedió como sigue: dada una ecuación general o particular, primero demostró cómo se encontraba el grupo G de esta ecuación en el cuerpo de los coeficientes, esto es, el grupo de las sustituciones de las raíces que deja invariante cada relación entre las raíces con los coeficientes en ese cuerpo. Por supuesto, se ha de encontrar el grupo de las ecuaciones sin conocer las raíces. En nuestro ejemplo anterior, el grupo de la ecuación cuártica era de orden 8 y el cuerpo de los coeficientes era R. Habiendo encontrado el grupo G de la ecuación, se busca en seguida el subgrupo mayor H en G. En nuestro ejemplo éste fue el subgrupo de orden 4. Si existieran dos o más subgrupos mayores, tomamos cualquiera. La determinación de H es una cuestión de pura teoría de grupos y es susceptible de realizarse. Habiendo encontrado H, se puede encontrar mediante un conjunto de procedimientos, considerando únicamente operaciones racionales, una función ϕ de las raíces cuyos coe-

ficientes pertenecen a R y que no cambia de valor bajo las sustituciones en H, pero sí varía con las otras sustituciones en G. En nuestro ejemplo anterior la función era $x_1^2 - x_3^2$. De hecho, es posible obtener una infinidad de tales funciones. Por supuesto, necesitamos encontrar tal función sin conocer las raíces. Existe un método de construcción de una ecuación en R una de cuyas raíces es la función ϕ. El grado de esta ecuación es el índice de H en G. Esta ecuación es llamada una resolvente parcial [12]. En nuestro ejemplo, la ecuación es $t^2 - (p^2 - 4q) = 0$ y su grado es 8/4 o bien 2.

Ahora bien, se debe ser capaz de resolver la resolvente parcial para encontrar la raíz ϕ. En nuestro ejemplo ϕ es $\sqrt{p^2 - 4q}$. Se adjunta ϕ a R obteniendo un nuevo cuerpo R'. Entonces el grupo de la ecuación original con respecto al cuerpo R' resulta ser H.

Ahora repetimos el procedimiento. Tenemos el grupo H, de orden 4 en nuestro ejemplo, y el cuerpo R', y buscamos el subgrupo mayor en H. En nuestro ejemplo, era el subgrupo de orden 2. Llamemos a este subgrupo K. Ahora obtenemos una función de las raíces de la ecuación original cuyos coeficientes pertenecen a R' y cuyo valor no varía bajo cada sustitución en K, pero es cambiado por otras sustituciones en H. En nuestro ejemplo esta ecuación es $t^2 - 2(-p - \sqrt{p^2 - 4q}) = 0$. El grado de dicha ecuación es el índice de K con respecto a H, esto es, 4/2 o 2. Tal ecuación es la segunda revolvente parcial.

Ahora se debe ser capaz de resolver esta ecuación revolvente y obtener una raíz, la función ϕ_1, y adjuntar este valor a R' y así formar el cuerpo R''. Con respecto a R'', el grupo de la ecuación es K.

Repetimos de nuevo el proceso. Encontramos el subgrupo máximo L en K. En nuestro ejemplo éste era justamente la sustitución identidad E. Buscamos una función de las raíces (con coeficientes en R'') que conserve su valor bajo E pero no por otras sustituciones de K. En nuestro ejemplo tal función era la $x_1 - x_2$. Para obtener esta ϕ_2 sin conocer las raíces debemos construir una ecuación en R'' teniendo la función ϕ_2 como raíz. En nuestro ejemplo la ecuación es $t^2 - 2(-p + \sqrt{p^2 - 4q}) = 0$. El grado de esta ecuación es el índice de L en K. En nuestro ejemplo este índice era 2/1 o 2. Tal ecuación es la tercera revolvente parcial. Debemos resolverla para encontrar el valor de ϕ_2.

Al adjuntar esta raíz a R'' obtenemos un cuerpo R'''. Suponga-

[12] Este uso de la palabra «resolvente» es diferente del de Lagrange.

mos que hemos llegado al estadio final donde el cuerpo de la ecuación original en R''' es la substitución identidad E.

Galois demostró que cuando el grupo de una ecuación con respecto a un cuerpo dado es justamente E, entonces las raíces de la ecuación son elementos de ese cuerpo. Por tanto las raíces están en el cuerpo R''' y conocemos el cuerpo en que están las raíces, porque R''' fue obtenido a partir del cuerpo conocido R por la adjunción de cantidades conocidas. Existe un proceso directo para encontrar las raíces mediante operaciones racionales en R'''.

Galois proporcionó un método para encontrar el grupo de una ecuación dada, las resolventes sucesivas y los grupos de la ecuación con respecto a los cuerpos de coeficientes sucesivamente ampliados que resultan de añadir las raíces de estas resolventes sucesivas del grupo original. Estos procesos incluyen mucha teoría pero, como señaló el propio Galois, su trabajo no estaba dirigido a ser un método práctico para resolver ecuaciones.

Luego, Galois aplicó la teoría anterior al problema de resolver ecuaciones polinomiales por medio de operaciones racionales y radicales. Aquí presentó otra noción de la teoría de grupos. Supongamos que H es un subgrupo de G. Si se multiplican las sustituciones de H por cualquier elemento g de G, entonces se obtiene una nueva colección de sustituciones que será denotada por gH, indicando la notación que la sustitución g primero se realiza y luego es aplicado cualquier elemento de H. Si $gH = Hg$ para cada g en G, entonces H es llamado un subgrupo normal (autoconjugado o invariante) en G.

Recordamos que el método de Galois de resolver una ecuación requiere encontrar y resolver las resolventes sucesivas. Galois demostró que cuando el resolvente que sirve para reducir el grupo de una ecuación, digamos de G a H, es una ecuación binómica $x^p = A$ de grado primo p, entonces H es un subgrupo normal en G (y de índice p), e inversamente, si H es un grupo subnormal en G y de índice primo p entonces la resolvente correspondiente es una ecuación binómica de grado p o puede ser reducida a una de ese tipo. Si todas las resolventes sucesivas son ecuaciones binomiales, podemos resolver la ecuación original mediante radicales, pues sabemos que podemos pasar del cuerpo inicial al cuerpo final en el que están las raíces mediante adjunciones sucesivas de radicales. Inversamente, si una ecuación es soluble por radicales, entonces el conjunto de las ecuaciones resolventes debe existir, y son ecuaciones binomiales.

Así, la teoría de la resolubilidad mediante radicales es, en una visión general, la misma que la teoría de la solución dada anteriormente, excepto que en la serie de subgrupos

$$G, H, K, L, ..., E$$

cada uno debe ser un subgrupo máximo normal (no ser un subgrupo de cualquier subgrupo normal mayor) del grupo precedente. Tal serie es llamada una serie de composición. Los índices de H en G, K en H y así en adelante, son llamados los índices de la serie de composición. Si los índices son números *primos* la ecuación es soluble por radicales, y si los índices no son primos no lo es. Cuando uno continúa hasta encontrar la sucesión de subgrupos normales maximales puede haber una elección; esto es, puede existir más de un subgrupo normal maximal de orden máximo en un grupo dado o subgrupo. Uno puede escoger cualquiera, aunque a partir de ahí los subgrupos pueden diferir. Pero resultará el mismo conjunto de índices aunque el orden en el que aparezcan difiera (véase el teorema de Jordan-Hölder más adelante). El grupo G, que contiene una serie de composición de índices primos, se dice que es resoluble.

¿Cómo es que la teoría de Galois muestra que la ecuación general de grado n no es soluble por radicales para $n > 4$ mientras que para $n \leqslant 4$ sí lo es? Para la ecuación general de grado n-ésimo el grupo está compuesto por todas las $n!$ substituciones de las n raíces. Este grupo es llamado el grupo simétrico de grado n. Su orden es por supuesto $n!$. No es difícil encontrar la serie de composición para cada grupo simétrico. El subgrupo normal maximal, que es llamado el subgrupo alternado, es de orden $n!/2$. El único subgrupo normal del grupo alternado es el elemento identidad. De aquí que los índices sean 2 y $n!/2$. Pero el número $n!/2$, para $n > 4$, nunca es primo. De aquí que la ecuación general de grado mayor que 4 no sea soluble por radicales. Por otra parte, la ecuación cuadrática puede ser resuelta con ayuda de una única ecuación resolvente. Los índices de la serie de composición consisten únicamente en el número 2. La ecuación general de tercer grado requiere para su solución dos ecuaciones resolventes de la forma $y^2 = A$ y $z^3 = B$. Estas son, por supuesto, resolventes binomiales. Los índices de la serie de composición son 2 y 3. La ecuación general de cuarto grado puede ser resuelta con cuatro ecuaciones resolventes binomiales, una de grado 3 y tres de grado 2. Los índices de la serie de composición son, entonces, 2, 3, 2, 2.

Para ecuaciones con coeficientes numéricos, como opuestas a aquellas con coeficientes literales independientes, Galois proporcionó una teoría similar a la descrita con anterioridad. Sin embargo, el proceso de determinar la solubilidad mediante radicales es más complicado, a pesar de que los principios básicos son los mismos.

Galois demostró también algunos teoremas especiales. Si uno tiene una ecuación irreducible de grado primo cuyos coeficientes están en un cuerpo R y cuyas raíces son todas funciones racionales de dos de la raíces con coeficientes en R, entonces la ecuación es soluble mediante radicales. También demostró el recíproco: toda ecuación irreducible de grado primo que es soluble mediante radicales tiene la propiedad que cada una de las raíces es una función racional de dos de ellas con coeficientes en R. Tal ecuación es llamada ahora ecuación galoisiana. El ejemplo más simple de ecuación galoisiana es $x^p - A = 0$. La noción es una extensión de las ecuaciones abelianas.

Como un epílogo, Hermite [13] and Kronecker (en una carta a Hermite) [14] y en un artículo posterior [15] resolvieron la ecuación de quinto grado por medio de funciones modulares elípticas. Esto es análogo al uso de funciones trigonométricas para resolver el caso irreducible de la ecuación cúbica.

5. Los problemas de construcción geométrica

Los matemáticos del siglo XVIII no dudaron que los famosos problemas de construcción pudieran ser resueltos. El trabajo de Galois proporcionó un criterio de constructibilidad que eliminó algunos de los famosos problemas.

Cada caso de una construcción con regla y compás requiere encontrar un punto de intersección, ya sea de dos rectas, una recta y un círculo, o dos círculos. Con la introducción de la geometría de coordenadas se reconoció que, en términos algebraicos, tales pasos significan la solución simultánea de dos ecuaciones lineales, una lineal o una cuadrática, o dos ecuaciones cuadráticas. En culquier caso, lo peor que interviene algebraicamente es una raíz cuadrada.

---

[13] *Comp. Rend.*, 46, 1858, 508-515 = *Œuvres*, 2, 5-12.
[14] *Comp. Rend.*, 46, 1858, 1150-1152 = *Werke*, 4, 43-48.
[15] *Jour. fur Math.*, 59, 1861, 306-310 = *Werke*, 4, 53-62.

De aquí que las cantidades sucesivas encontradas por pasos o construcciones sucesivas son en el peor de los casos el resultado de una cadena de *raíces cuadradas* que se aplican a cantidades dadas. Por consiguiente, las cantidades constructibles deben estar en cuerpos obtenidos por la adjunción al cuerpo conteniendo las cantidades dadas de únicamente raíces cuadradas de cantidades dadas o construidas después. Podemos llamar a tales extensiones cuerpos de extensión cuadrática.

Al llevar a cabo las construcciones sucesivas deben ser observadas unas cuantas restricciones. Por ejemplo, algunos de los procesos permiten el uso de una recta o círculo arbitrario. Así, al bisecar un segmento de recta podemos usar círculos más grandes que la mitad del segmento de recta. Uno debe escoger este círculo en el cuerpo o en el cuerpo de extensión construible de los elementos dados. Esto puede ser hecho.

También sucede que los cuerpos de extensión pueden contener elementos complejos porque, por ejemplo, puede aparecer la raíz cuadrada de una coordenada negativa. Estos elementos complejos son constructibles, porque las partes real e imaginaria de las cantidades complejas que aparecen son cada una de ellas raíces de una ecuación real y estas raíces son construibles.

Dado un problema de construcción, uno halla primero una ecuación algebraica cuya solución es la cantidad deseada. Esta cantidad debe pertenecer a algún cuerpo de extensión cuadrática del cuerpo de las cantidades dadas. En el caso del polígono regular de 17 lados esta ecuación es $x^{17} - 1 = 0$ y la cantidad dada puede ser tomada como el radio del círculo unitario. La ecuación irreducible relevante es $x^{16} + x^{15} + ... + 1 = 0$. En términos de la teoría de Galois la condición necesaria y suficiente para que una ecuación sea soluble con *raíces cuadradas* es que el orden del grupo de Galois de la ecuación sea una potencia de 2. Este es el caso para la ecuación $x^{16} + ... + 1 = 0$ y la serie de composición es 2,2,2,2. Esto significa que las resolventes son ecuaciones binomiales de grado 2, de tal forma que únicamente son añadidas raíces cuadradas al cuerpo racional original determinado por el radio dado del círculo unidad. Con este criterio de Galois se puede demostrar la proposición de Gauss de que un polígono regular con un número p primo de lados puede ser construido con regla y compás si y sólo si el número p tiene la forma $2^{2^n} + 1$, esto es, si $p = 3, 5, 17, 257, ...$ pero no para $p = 7, 11, 13, 19, 23, 29, 31, ...$ La teoría de Galois puede ser utili-

zada para demostrar también que es imposible trisecar un ángulo arbitrario o duplicar un cubo dado.

Pero el criterio de Galois no se aplica en absoluto al problema de cuadrar el círculo. Aquí la cantidad dada es el radio del círculo. La ecuación correspondiente es $x^2 = \pi r^2$. A pesar de que esta ecuación es únicamente cuadrática, no es cierto que su solución pertenece a un cuerpo de extensión cuadrática del cuerpo determinado por la cantidad dada, ya que π no es un irracional algebraico (cap. 41, sec. 2). El trabajo de Galois, entonces, no sólo respondió por completo a la cuestión de qué ecuaciones son solubles con operaciones algebraicas, sino que dio un criterio general para determinar la constructibilidad con regla y compás de figuras geométricas.

En cuanto concierne a los famosos problemas de construcción, debemos decir que antes de que la teoría de Galois se les aplicara, Gauss y Wantzel habían determinado cuáles de los polígonos regulares son construibles (sec. 2), y Wantzel, en un ensayo de 1837 [16], demostró que el ángulo general no podía ser trisecado, ni tampoco se podía duplicar un cubo. Demostró que cada cantidad constructible debe satisfacer una ecuación de grado 2^n, y esto no es cierto en los dos casos mencionados.

6. La teoría de los grupos de sustituciones

En su trabajo sobre la solubilidad de ecuaciones (cap. 25, sec. 2), Lagrange presentó como la clave para su análisis las funciones de las n raíces que toman el mismo valor bajo algunas permutaciones de las raíces. A continuación, se propuso estudiar, en los mismos ensayos, funciones con el propósito de determinar los diferentes valores que toman con los $n!$ posibles valores que las $n!$ permutaciones de las n variables (las raíces) pueden originar. Trabajos posteriores de Ruffini, Abel y Galois prestaron una mayor importancia a este tema. El hecho de que una función racional de n letras tome el mismo valor bajo alguna colección de permutaciones o sustituciones de las raíces significa, como ya hemos visto, que esta colección es un subgrupo del grupo simétrico total. Esto fue observado explícitamente por Ruffini en su *Teoria generale delle equazioni* (*Teoría General de las Ecuaciones*, 1799). De aquí que lo que Lagrange inició

[16] *Jour. de Math.*, 2, 1837, 366-372.

significara una manera de estudiar subgrupos de un grupo de sustituciones. El modo más directo es, por supuesto, estudiar el propio grupo de sustituciones y determinar sus subgrupos. Ambos métodos para estudiar la estructura o composición de grupos de sustituciones se convirtieron en una materia activa seguida con un interés en sí y para sí misma, a pesar que no se ignoró la conexión con la solubilidad de las ecuaciones. La teoría de sustituciones o grupos de permutaciones fue la primera gran investigación que en última instancia promovió el surgimiento de la teoría abstracta de grupos. Aquí debemos hacer mención de algunos teoremas concretos sobre grupos de sustituciones que se obtuvieron durante el siglo XIX.

El propio Lagrange estableció un resultado importante, que en lenguaje moderno afirma que el orden de un subgrupo divide el orden del grupo. Pietro Abbati (1768-1842), en una carta del 30 de septiembre de 1802 —publicada posteriormente— [17] comunicó a Ruffini la demostración de este teorema.

En su libro de 1799 Ruffini presentó, aunque de una manera vaga, las nociones de transitividad y primitividad. Un grupo de permutaciones es transitivo si cada letra del grupo es reemplazada por cada una de las otras letras bajo las varias permutaciones del grupo. Si G es un grupo transitivo y los n símbolos o letras pueden ser divididos en r subconjuntos diferentes σ_i, $i = 1, 2, \ldots, r$, conteniendo cada subconjunto s_i símbolos de tal forma que cualquier permutación de G o permuta los simbolos de las σ_i entre ellos mismos, o reemplaza estos símbolos por los símbolos de las σ_j, esto para cada $i = 1,2,\ldots r$, entonces G es llamado imprimitivo. Si no es posible tal separación de los n símbolos, entonces el grupo transitivo es llamado grupo primitivo. Ruffini demostró también que no existe un subgrupo de orden k para todo k en un grupo de orden n.

Cauchy, motivado por el trabajo de Lagrange y Ruffini, escribió un artículo importante sobre los grupos de sustituciones [18]. Con la teoría de ecuaciones a la vista, demostró que no existe un grupo de n letras (grado n) cuyo índice relativo a la totalidad del grupo simétrico en n letras sea menor que el máximo número primo que no excede n, a menos que el índice sea 2 ó 1. Cauchy estableció este teorema en el lenguaje de los valores de las funciones: el número de valores diferentes de una función no simétrica de n letras no puede

[17] *Memoire della Società Italiana delle Scienze*, 10, 1803, 385-409.
[18] *Jour. de l'Ecole Poly.*, 10, 1815, 1-28 = *Œuvres* (2), 1, 64-90.

ser menor que el máximo primo p menor que n, a menos que sea 2.

Galois realizó el mayor avance en la introducción de conceptos y teoremas acerca de grupos de sustituciones. Su concepto más importante lo constituyó la nocion de subgrupo normal (invariante o autoconjugado). Otro concepto para grupos debido a Galois es el de isomorfismo entre dos grupos, esto es, una correspondencia uno-a-uno entre los elementos de los dos grupos tal que si $a.b. = c$ en el primero, entonces para los elementos correspondientes en el segundo $a'.b' = c'$. También presentó las nociones de grupos simples y compuestos. Un grupo que carece de subgrupo invariante es simple; si no es compuesto. A propósito de estas nociones, Galois expresó la conjetura [19] de que el grupo simple más pequeño cuyo orden es un número compuesto es un grupo de orden 60.

El trabajo de Galois no fue conocido hasta que Liouville publicó partes de él en 1846, y aun entonces el contenido no era fácilmente accesible. Por otro lado, los trabajos de Lagrange y Ruffini sobre grupos de sustituciones, apoyados en el lenguaje de los valores que puede tomar una función de n letras, se hicieron bien conocidos. De aquí que la solución de ecuaciones retrocediera, y cuando Cauchy volvió a la teoría de las ecuaciones se concentró en los grupos de sustituciones. Durante los años de 1844 a 1846 escribió multitud de ensayos. En el principal [20], sistematizó muchos de los resultados anteriores y demostró un buen número de teoremas especiales sobre grupos primitivos, transitivos y demostró la aserción de Galois de que todo grupo finito (de sustituciones) cuyo orden es divisible por un primo p contiene al menos un subgrupo de orden p. El artículo principal fue seguido por un gran número de otros publicados en las *Comptes Rendus* de la Academia de París entre los años 1844 y 1846 [21]. La mayor parte de su trabajo versaba sobre valores formales (esto es no numéricos) que las funciones de n letras pueden tomar al cambiarse las letras y en encontrar funciones que toman un número dado de valores.

Después de que Liouville publicase parte del trabajo de Galois, Serret impartió cursos sobre ello en la Sorbona y en la tercera edición de su *Cours* proporcionó una mejor exposición textual de la

[19] *Œuvres*, 1897, ed., 26.

[20] *Exercices d'analyse et de physique mathématique*, 3, 1844, 151-252 = *Œuvres* (2), 13, 171-282.

[21] *Œuvres* (1), vols. 9 y 10.

teoría de Galois. El trabajo de clarificar las ideas de Galois sobre la solubilidad de ecuaciones y el desarrollo de la teoría de los grupos de sustituciones continuaron a la par a partir de ahí. En su texto, Serret proporcionó una forma mejorada del resultado de Cauchy de 1815. Si una función de n letras toma menos de p valores, donde p es el mayor primo menor que n, entonces la función no admite más de dos valores.

Uno de los problemas que Serret subrayó en el texto de 1866 pide todos los grupos que pueden formarse con n letras. Este problema ya había atraído la atención de Ruffini; y él, Cauchy y el mismo Serret, en su ensayo de 1850 [22], proporcionaron un buen número de resultados parciales, como también lo hizo Thomas Penyngton Kirman (1806-1895). A pesar de muchos esfuerzos y cientos de resultados parciales, el problema sigue sin solución.

Después de Galois, Camile Jordan (1838-1922) fue el primero en añadir conceptos significativos a la teoría de Galois. En 1869 [23], demostró un resultado básico. Sea G_1 un subgrupo maximal autoconjugado (normal) de G_0, G_2 un subgrupo maximal autoconjugado de G_1, y así en adelante hasta que la serie termina en el elemento identidad. Esta serie de subgrupos es llamada una serie de composición de G_0. Si G_{i+1} es cualquier subgrupo autoconjugado de orden r en G_i cuyo orden es p, entonces G_i puede ser descompuesto en $\lambda = p/r$ clases. Dos elementos están en la misma clase si uno es el producto del otro y un elemento de G_{i+1}. Si a es cualquier elemento en una clase y b cualquier elemento en otra, el producto estará en la misma tercera clase. Estas clases forman un grupo para el cual G_{i+1} es el elemento identidad y el grupo es llamado grupo cociente o grupo factor de G_i por G_{i+1}. Se denota por G_i/G_{i+1}, notación introducida por Jordan en 1872. Los grupos cociente G_0/G_1, G_1/G_2, ... son llamados los grupos factores de composición de G_0 y sus órdenes se conocen como los factores de composición o índices de composición. Puede haber más de una serie de composición en G_0. Jordan demostró que el conjunto de factores de composición es invariante excepto para el orden en el que pueden aparecer y (Ludwig) Otto Hölder (1859-1937), profesor de la Universidad de Leipzig, demostró [24] que los grupos cocientes mismos eran independientes

[22] *Jour. de Math.* (1), 15, 1850, 45-70.
[23] *Jour de Math.* (2), 14, 1869, 129-146 = *Œuvres,* 1, 241-248.
[24] *Math. Ann.*, 34, 1889, 26-56.

de las series de composición; esto es, el mismo conjunto de grupos cocientes estaría presente para cualquier serie de composición. Los dos resultados son llamados el teorema de Jordan-Hölder.

El conocimiento de los grupos de sustituciones (finitos) y su conexión con la teoría de ecuaciones de Galois hasta 1870, fue organizado en un libro magistral de Jordan, su *Traité des substitutions et des équations algébriques* (*Tratado de las sustituciones y de las ecuaciones algebraicas,* 1870). En su libro, Jordan —como casi todos los autores que lo precedieron— usó como definición de grupo de sustituciones la de que es una colección de sustituciones tales que el producto de dos miembros cualesquiera de la colección pertenece a la colección. Las otras propiedades que comúnmente postulamos hoy en día en la definición de un grupo (cap. 49, sec. 2) se utilizaron, pero fueron incorporadas como propiedades obvias de tales grupos o como condiciones adicionales, aunque no especificadas en la definición. El *Traité* presentó nuevos resultados e hizo explícitos para grupos de sustituciones las nociones de isomorfismo *(isomorphisme holoédrique)* y homomorfismo *(isomorphisme mériédrique),* siendo la última una correspondencia de *varios-a-uno* entre los dos grupos tal que $a.b = c$ implica $a'.b' = c'$. Jordan añadió resultados fundamentales sobre grupos transitivos y compuestos. El libro también contiene la solución de Jordan al problema propuesto por Abel, determinar las ecuaciones de un grado dado que son resolubles por radicales y reconocer si una ecuación dada pertenece o no pertenece a esta clase. Los grupos de las ecuaciones solubles son conmutativos. Jordan los llamó abelianos y el término abeliano se aplicó a los grupos conmutativos a partir de entonces.

Poco tiempo después de haber aparecido el *Traité,* el profesor de matemáticas noruego Ludwig Sylow (1832-1918) demostró otro teorema importante sobre grupos de sustituciones. Cauchy había demostrado que todo grupo cuyo orden es divisible por un número primo p debe contener uno o más subgrupos de orden p. Sylow[25] amplió el teorema de Cauchy. Si el orden de un grupo es divisible por p, siendo p primo, pero no por $p^{\alpha+1}$, entonces el grupo contiene uno y sólo un sistema de subgrupos conjugados de orden p^{α} [26].

[25] *Math. Ann.*, 5, 1872, 584-594.

[26] Si H es un subgrupo de G y q es cualquier elemento de G, entonces $q^{-1}Hq$ es un subgrupo conjugado de H. Se dice que H y todos sus conjugados forman un sistema de subgrupos conjugados de G, o un conjunto completo conjugado de subgrupos.

Sylow prueba en el mismo artículo que todo grupo de orden p^α es resoluble, esto es, que los índices de una sucesión de subgrupos maximales invariantes son primos.

Otro enfoque diferente de los grupos de sustituciones, y en última instancia de los grupos generales, fue sugerido por una investigación puramente física. Auguste Bravais (1811-1863), físico y mineralólogo, estudió grupos de movimientos [27] para determinar las estructuras posibles de los cristales. Este estudio se reduce matemáticamente a la investigación de las transformaciones lineales en tres variables

$$x_i' = a_{i1}x + a_{i2}y + a_{i3}z, \qquad i = 1, 2, 3,$$

de determinante +1 o −1 y condujo a Bravais a treinta y dos clases de estructuras moleculares simétricas que pueden aparecer en los cristales.

El trabajo de Bravais impresionó a Jordan y se propuso investigar lo que llamó la representación analítica de grupos y que ahora se denomina la teoría de representación de grupos. De hecho, Serret, en la edición de 1866 de su *Cours*, había considerado la representación de sustituciones mediante transformaciones de la forma

$$x' = \frac{ax + b}{cx + d}.$$

Pero la representación más útil de todos los tipos de grupos la introdujo Jordan. Buscó representar sustituciones mediante transformaciones lineales de la forma

$$x_i' = \sum_{j=1}^{n} a_{ij}x_j, \qquad i = 1, 2, ..., n. \tag{9}$$

Ya que los grupos de sustituciones son finitos, tienen que imponerse algunas restricciones a las transformaciones de tal manera que el grupo de transformaciones sea finito. Galois había considerado tales transformaciones [28] y las limitó de tal forma que los coeficientes y las variables toman valores en un cuerpo finito de orden primo. En

[27] *Jour. de Math.*, 14, 1849, 141-180.
[28] *Œuvres*, 1897, ed., 21-23, 27-29.

1878 [29], Jordan enunció que una sustitución homogénea lineal (9) de período finito p puede ser transformada linealmente en la forma canónica

$$y_i' = \varepsilon_i y_i, \qquad i = 1, 2, \ldots, n$$

donde las $\varepsilon_{i's}$ son raíces p-ésimas de la unidad. El teorema lo demostraron varios autores [30], y esto constituyó el inicio de buen número de investigaciones sobre la determinación de los posibles grupos de sustituciones lineales de orden dado y de una forma binaria o ternaria (dos y tres variables). También la determinación de los subgrupos de grupos de sustituciones lineales y las expresiones algebraicas invariantes por todos los miembros de un grupo o subgrupo promovieron mucha investigación.

Inmediatamente después de conocer el ensayo de Bravais, Jordan se propuso la primera gran investigación en grupos *infinitos*. En su ensayo «Mémoire sur les groupes de mouvements» («Memoria sobre los grupos de movimientos») [31], Jordan señala que la determinación de todos los grupos de movimientos (únicamente consideraba traslaciones y rotaciones) es equivalente a la determinación de todos los posibles sistemas de moléculas tales que cada movimiento de cualquier grupo transforma el sistema correspondiente de moléculas en sí mismo. Por tanto, estudió los varios tipos de grupos y los clasificó. Los resultados no son tan significativos como el hecho de que su artículo inició el estudio de las transformaciones geométricas en el marco de los grupos y los geómetras se apresuraron a seguir esta línea de pensamiento (cap. 38, sec. 5).

Otro desarrollo de la mitad del siglo XIX es tan notable como instructivo. Arthur Cayley, muy influenciado por el trabajo de Cauchy, reconoció que la noción de grupo de sustituciones podía ser generalizada. En tres artículos [32], Cayley presentó la noción de un grupo *abstracto*. Utilizó un símbolo de operador general θ aplicado a un sistema de elementos $x,y,z,\ldots$ y habló de θ como aplicado para dar una función $x',y',z',\ldots$ de $x,y,z,\ldots$ Señaló que en particular

[29] *Jour. für Math.*, 84, 1878, 89-215, p. 112 en particular = *Œuvres*, 2, 13-139, p. 36 en particular.

[30] Véase: E. H. Moore, *Math. Ann.*, 50, 1898, 215.

[31] *Annali di Mat.* (2), 2, 1868/1869, 167-215 y 322-345 = *Œuvres*, 4, 231-302.

[32] *Phil. Mag.* (3), 1849, 527-529 = *Coll. Math. Papers*, 1, 423-424 y (4), 7, 1854, 40-47 y 408-409 = *Papers*, 2, 123-130 y 131-132.

θ puede ser una sustitución. El grupo abstracto contiene muchos operadores θ, $\varnothing$,.... $\theta\varnothing$ es un (producto) compuesto de dos operaciones y el compuesto es asociativo pero no necesariamente conmutativo. Su definición general de grupo supone un conjunto de operadores 1, α, p,..., todos ellos diferentes y tales que el producto de dos cualesquiera de ellos en cualquier orden, o el producto de cualquiera por sí mismo, pertenece al conjunto [33]. Menciona las matrices bajo la multiplicación y los cuaterniones (con la adición) como constituyendo grupos. Desafortunadamente, la introducción por Cayley del concepto de grupo abstracto no atrajo la atención en su tiempo, en parte porque las matrices y los cuaterniones eran nuevos y no bien conocidos, y los otros muchos sistemas matemáticos que podían ser incluidos en la noción de grupo o estaban aún por desarrollarse o no se había reconocido que lo eran. La abstracción prematura cae en oídos sordos, ya pertenezcan a matemáticos, ya a estudiantes.

Bibliografía

Abel, H. H.: *Œuvres complètes.* (1881), 2 vols., Johnson Reprint Corp., 1964.

Bachmann, P.: «Über Gauss' zahlentheoretische Arbeiten», *Nachrichten König. Ges. der. Wiss. zu Gött.,* 1911, 455-518. También en los *Werke* de Gauss, 10, Parte 2, 1-69.

Burkhardt, H.: «Endliche discrete Gruppen», *Encyk. der. Math. Wiss.,* B. G. Teubner, 1903-1915, I, Parte I, 208-226.

— : «Die Anfänge der Gruppentheorie und Paolo Ruffini», *Abhandlungen zur Geschichte der Mathematik,* Heft 6, 1892, 119-159.

Burns, Josephine E.: «The Foundation Period in the History of Group Theory», *Amer. Math. Monthly,* 20, 1913, 141-148.

Dupuy, P.: «La vie d'Evariste Galois», *Ann. de l'Ecole Norm. Sup.* (2), 13, 1896, 197-266.

Galois, Evariste: «Œuvres», *Jour. de Math.,* 11, 1846, 381-444.

— : *Oeuvres Mathématiques,* Gauthier-Villars, 1897.

— : *Ecrits et mémoires mathématiques* (ed. por R. Bourgne y J. P. Azra), Gauthier-Villars, 1962.

Gauss, C. F.: *Disquisitiones Arithmeticae* (1801), *Werke,* vol. 1, König. Ges. der Wiss., zu Göttingen, 1870. Traducción al inglés de Arthur A. Clarke, S. J., Yale University Press, 1966.

---

[33] *Papers,* 2, 124.

Hobson, E. W.: *Squaring the circle and other monographs*, Chelsea (reimpresión), 1953.
Hölder, Otto: «Galois'sche Theorie mit Anwendungen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1898-1904, I, Parte I, 480-520. Hay versión española, L. Infeld, *El elegido de los dioses,* México, Siglo XXI, 1974.
Infeld, Leopold: *El elegido de los dioses*. México, Siglo XXI, 1974.
Jordan, Camille: *Œuvres*, 4 vols., Gauthier-Villars, 1961-1964.
— : *Traité des substitutions et des équations algébriques* (1870), Gauthier-Villars (reimpresión), 1957.
Kiernan, B. M.: «The Development of Galois Theory from Lagrange to Artin», *Archive for History of Exact Sciences,* 8, 1971, 40-154.
Lebesgue, Henri: *Notice sur la vie et les travaux de Camille Jordan*, Gauthier-Villars, 1923. También en *Notices d'historie des mathématiques* de Lebesgue, pp. 44-65, Institut de Mathématiques, Genève, 1958.
Miller, G. A.: «History of the Theory of Groups to 1900», *Collected Works,* vol. 1, 427-467, University of Illinois Press, 1935.
Pierpont, James: «Lagrange's place in the Theory of Substitutions», *Amer. Math. Soc. Bull.,* 1, 1894-1895, 196-204.
— : «Early History of Galois Theory of Equations», *Amer. Math. Soc. Bull.,* 4, 1898, 332-340.
Smith, David Eugene: *A Source Book in Mathematics*, Dover (reimpresión), 1959, vol. 1, 232-252, 253-260, 261-266, 278-285.
Verriest, G.: *Œuvres mathématiques d'Evariste Galois* (ed. 1897), 2.ª ed., Gauthier-Villars, 1951.
Wiman, A.: «Endliche Gruppen linearer Substitutionen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1898-1904, I, Parte 1, 522-554.
Wussing, H. L.: *Die Genesis des abstrakten Gruppenbegriffes,* VEB Deutscher Verlag des Wiss., 1969.

Capítulo 32
CUATERNIONES, VECTORES Y ALGEBRAS LINEALES ASOCIATIVAS

> Los cuaterniones fueron descubiertos por Hamilton después de haber realizado su obra realmente importante y, aunque ingeniosamente bellos, han sido un demonio ambiguo para aquellos que los han tocado de alguna manera... El vector es un superviviente inútil, o una derivación de los cuaternios y nunca ha sido de la menor utilidad para criatura alguna.
>
> LORD KELVIN

1. El fundamento del álgebra sobre la permanencia de la forma

El trabajo de Galois sobre la solubilidad de ecuaciones mediante procesos algebraicos cerró un capítulo del álgebra y, a pesar de que presentó ideas tales como las de grupo y dominio de racionalidad (cuerpo) que rendirían fruto más tarde, la completa explotación de estas ideas tuvo que esperar otros desarrollos. La siguiente gran creación algebraica, iniciada por William R. Hamilton, abrió nuevos dominios, mientras rompía con viejas convicciones acerca de cómo debían comportarse los «números».

Para apreciar la originalidad del trabajo de Hamilton hay que examinar cómo se entendía la lógica del álgebra ordinaria en la primera parte del siglo XIX. Hacia 1800 los matemáticos empleaban con libertad los varios tipos de números reales y aun complejos, pero la definición precisa de estos distintos tipos de números no se hallaba a disposición de los matemáticos, ni tampoco existía una justificación de las operaciones con ellos. Las expresiones de insatisfacción hacia el estado de cosas fueron muy numerosas, pero estaban sumergidas en la masa de nuevas creaciones del álgebra y el análisis. Las mayores inquietudes parecían estar causadas por el hecho de que se

manipulaban las letras como si tuvieran las propiedades de los enteros; sin embargo, los resultados de estas operaciones eran válidos cuando números cualesquiera sustituían las letras. Ya que no se había realizado el desarrollo de la lógica de los diversos tipos de números, no era posible ver que éstos poseían las mismas propiedades formales de los enteros positivos y, consecuentemente, que expresiones literales que simplemente se mantenían para cualquier clase de números reales o complejos deben poseer las mismas propiedades —esto es, que el álgebra ordinaria es únicamente aritmética generalizada—. Parecía como si el álgebra de expresiones literales poseyera una lógica en sí misma, que respondía de su efectividad y corrección. De aquí que los matemáticos atacaran hacia 1830 el problema de justificar las operaciones con expresiones literales o simbólicas.

Este problema lo consideró primero George Peacock (1791-1858), profesor de matemáticas en la universidad de Cambridge. Para justificar las operaciones con expresiones literales que podían mantenerse para números negativos, irracionales y complejos hizo la distinción entre álgebra aritmética y álgebra simbólica. La primera trataba con símbolos representando los enteros positivos, y por lo mismo se encontraba sobre tierra firme. Aquí se permitían únicamente operaciones que condujeran a enteros positivos. El álgebra simbólica adopta las reglas del álgebra aritmética pero suprime las restricciones a enteros positivos. Todos los resultados deducidos en el álgebra aritmética, cuyas expresiones son generales en forma pero particulares en valor, son resultados de la misma manera que en álgebra simbólica, donde son generales en valor así como en forma. Así $a^m a^n = a^{m+n}$ se mantiene en álgebra aritmética cuando m y n son enteros positivos y por lo mismo se mantiene en álgebra simbólica para todas las m y n. Asimismo, la serie para $(a + b)^n$ donde n es un entero positivo, si fuera presentada en forma general sin referencia a un término final, es válida para todas las n. El argumento de Peacock es conocido como el principio de permanencia de la forma.

La formulación explícita de este principio fue proporcionada en el «Report on the Recent Progress and Present State of Certain Branches of Analysis» («Informe sobre los avances recientes y estado presente de ciertas ramas del análisis») de Peacock [1], en el que no sólo informa sino que afirma dogmáticamente. En el álgebra simbólica, dice:

---

[1] *Brit. Assn. for Adv. of Sci.*, Rept. 3, 185-352.

1. Los símbolos son ilimitados tanto en valor como en representación.
2. Las operaciones sobre ellos, cualesquiera que sean, son posibles en todos los casos.
3. Las leyes de combinación de los símbolos son de tal clase que coinciden universalmente con las del álgebra aritmética cuando estos símbolos son cantidades aritméticas, y cuando las operaciones a las que se sujetan son llamadas con los mismos nombres que en el álgebra aritmética.

A partir de estos principios creyó que era posible deducir el principio de permanencia de la forma: «Cualesquiera formas algebraicas que son equivalentes cuando los símbolos son generales en forma pero específicos en valores [enteros positivos], serán equivalentes de la misma manera cuando los símbolos son generales tanto en valor como en forma.» Peacock usó este principio para justificar en particular las operaciones con números complejos. Trató de proteger su conclusión mediante la frase «cuando los símbolos son generales en forma». Así se podía no establecer propiedades especiales de números enteros particulares en forma simbólica e insistir en que estos enunciados simbólicos son generales. Por ejemplo, la descomposición de un entero compuesto en producto de primos, aunque expresado simbólicamente, no era válido como enunciado del álgebra simbólica. El principio sancionaba por decreto lo que era correcto y evidentemente empírico, pero aún no establecido lógicamente.

Peacock reafirmó este principio en la segunda edición de su *Treatise on Algebra (Tratado de Algebra)*[2], pero aquí también presenta una ciencia formal del álgebra. En este *Tratado,* Peacock establece que el álgebra, como la geometría, es una ciencia deductiva. Los procesos del álgebra tienen que estar basados sobre un enunciado completo del cuerpo de las leyes que dictan las operaciones usadas en el proceso. Los símbolos para las operaciones no tienen, al menos para la ciencia deductiva del álgebra, ningún otro sentido que aquel dado por las leyes. Así, la adición no significa más que cualquier proceso que obedece las leyes de la adición del álgebra. Sus leyes son, por ejemplo, las leyes asociativa y conmutativa de la adición y la multiplicación, y la ley de que si $ac = bc$ y $c \neq 0$, entonces $a = b$.

Aquí fue derivado el principio de la permanencia de la forma de la adopción de los axiomas. Este enfoque allanó el camino para un

[2] 1842-1845; 1.ª ed., 1830.

pensamiento más abstracto en el álgebra y en particular influenció el pensamiento de Boole sobre el álgebra de la lógica.

A lo largo de la mayor parte del siglo XIX, se aceptó la visión del álgebra afirmada por Peacock. Fue apoyada, por ejemplo, por Duncan F. Gregory (1813-1844), un tataratataranieto de James Gregory, del siglo XVII. Gregory escribió en un ensayo «On the real nature of symbolic algebra» («Sobre la verdadera naturaleza del álgebra simbólica») [3]:

> La luz a la que consideraré el álgebra simbólica es como la ciencia que trata la combinación de las operaciones definidas, no por su naturaleza, esto es, por lo que son o lo que hacen, sino por las leyes de las combinaciones a las que están sujetas... Es cierto que estas leyes han sido en muchos casos sugeridas (como Mr. Peacock ha dicho convenientemente) por las leyes conocidas de las operaciones de los números, pero el paso dado del álgebra aritmética a la simbólica es que, dejando de lado la naturaleza de las operaciones que los símbolos que usamos representan, suponemos la existencia de clases de operaciones desconocidas sujetas a estas mismas leyes. Así somos capaces de probar ciertas relaciones entre las diferentes clases de operaciones, que, cuando son expresadas entre los símbolos, se llaman teoremas algebraicos.

En este ensayo, Gregory insistió en las leyes conmutativa y distributiva, términos introducidos por François-Joseph Servois (1767-1847) [4].

La teoría del álgebra como ciencia de los símbolos y las leyes de sus combinaciones fue llevada más lejos por Augustus de Morgan, quien escribiera varios artículos sobre la estructura del álgebra [5]. Su *Trigonometry and Double Algebra* (*Trigonometría y Algebra Doble*, 1849) contiene también sus puntos de vista. Las palabras álgebra doble significaban el álgebra de los números complejos, mientras que álgebra simple significaba el álgebra de los números negativos. Anterior al álgebra simple es la aritmética universal, que cubre el álgebra de los números reales positivos. El álgebra —decía de Morgan— es una colección de símbolos carentes de significado y operaciones entre estos símbolos. Los símbolos son 0, 1, +, −, ×, ÷, $()^{()}$, y letras. Las leyes del álgebra son las leyes que obedecen estos sím-

[3] *Transactions of the Royal Society of Edinburgh*, 14, 1840, 208-216.
[4] *Ann. de Math.*, 5, 1814-1815, 93-140.
[5] *Trans. Camb. Phil. Soc.*, 1841, 1842, 1844 y 1847.

bolos, por ejemplo, la ley conmutativa, la ley distributiva, las leyes de los exponentes, un número negativo por un positivo es negativo, $a - a = 0$, $a \div a = 1$, y las leyes derivadas. Las leyes básicas son seleccionadas arbitrariamente.

Los axiomas del álgebra aceptados a mediados del siglo XIX son:

1. Cantidades iguales sumadas a una tercera dan cantidades iguales.
2. $(a + b) + c = a + (b + c)$.
3. $a + b = b + a$.
4. Iguales añadidos a iguales dan iguales.
5. Iguales añadidos a desiguales dan desiguales.
6. $a(bc) = (ab)c$
7. $ab = ba$.
8. $a(b + c) = ab + ac$.

El principio de permanencia de la forma se apoyaba sobre estos axiomas.

Es difícil para nosotros ver lo que este principio significa exactamente. Formula la cuestión de por qué los varios tipos de números poseen las mismas propiedades que los números enteros. Pero Peacock, Gregory y De Morgan querían hacer del álgebra una ciencia independiente de las propiedades de los números reales y complejos y por lo mismo veían el álgebra como una ciencia de símbolos sin interpretación y sus leyes de combinación. En efecto, fue ésa la justificación de la suposición que las mismas propiedades fundamentales se mantenían para todos los tipos de números. Esta fundamentación no solamente era vaga, sino también poco elástica. Los autores insistían en un paralelismo entre el álgebra aritmética y la general tan rígido que, si se mantenía, destruiría la generalidad del álgebra. Parece que no se percataban de que una fórmula que es verdadera con una interpretacion de los símbolos puede no serlo con otra.

El principio de la permanencia de la forma, un dictado arbitrario, no podía servir como fundamento sólido para el álgebra. De hecho, los desarrollos con los que trataremos en este capítulo lo desacreditan. El primer paso, que hizo obvia la necesidad para los números complejos de este principio, lo dio Hamilton cuando fundó la lógica de los números complejos sobre las propiedades de los números reales.

Hamilton no estaba satisfecho con un fundamento meramente intuitivo refiriéndose a los números complejos, aunque éstos estu-

vieran fundados intuitivamente, al representarlos como puntos o segmentos de recta dirigidos en el plano. En su artículo «Conjugate Functions and on Algebra as the Sciencie of Pure Time» («Funciones conjugadas y sobre el álgebra como la ciencia del tiempo puro»)[6], Hamilton señaló que un número complejo $a + bi$ no es una suma en el sentido en que $2 + 3$ lo es. El uso del signo más es un accidente histórico y bi no puede ser añadido a a. El número complejo $a + bi$ no es más que una pareja ordenada (a,b) de números reales. La peculiaridad que i ó $\sqrt{-1}$ introduce en las operaciones con los números complejos está incorporada por Hamilton en las definiciones de las operaciones con pares ordenados. Así, si $a + bi$ y $c + di$ son dos números complejos entonces

$$\begin{aligned} (a, b) \pm (c, d) &= (a \pm c, b \pm d) \\ (a, b) \cdot (c, d) &= (ac - bd, ad + bc) \\ \frac{(a, b)}{(c, d)} &= \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2}\right). \end{aligned} \tag{1}$$

Las propiedades usuales asociativa, conmutativa y distributiva pueden ser deducidas ahora. Con esta visión de los números complejos, no sólo están estos números fundamentados lógicamente sobre la base de los números reales, sino que también el misterioso $\sqrt{-1}$ es evitado por completo. Por supuesto, en la práctica todavía es conveniente valerse de la forma $a + bi$ y recordar que $\sqrt{-1}\,\sqrt{-1} = -1$. Incidentalmente, Gauss dijo en una carta de 1837 a Wolfgang Bolyai que había tenido esta noción de las parejas ordenadas desde 1831. Pero fue la publicación de Hamilton la que proporcionó al mundo matemático la noción de par ordenado.

2. La búsqueda de un «número complejo» tridimensional

La noción de vector, esto es, un segmento de línea dirigido que puede representar la magnitud y dirección de una fuerza, una velocidad o una aceleración, entró en las matemáticas calladamente. Aristóteles sabía que las fuerzas podían ser representadas como vectores y que la acción combinada de dos fuerzas se obtenía por lo que es

[6] *Trans. Royal Irish Academy*, 17, 1837, 293-422 = *Math. Papers*, 3, 3-96.

vulgarmente conocido como la ley del paralelogramo; esto es, que la diagonal del paralelogramo formado por los dos vectores *a* y *b* (Fig. 32.1), proporciona la magnitud y dirección de la fuerza resultante. Simon Stevin empleó la ley del paralelogramo en problemas de estática y Galileo enunció la ley explícitamente.

Después de que la representación geométrica de los números complejos proporcionada por Wessel, Argand y Gauss fuera algo familiar, los matemáticos se percataron de que los números complejos podían usarse para trabajar y representar los vectores en un plano. Por ejemplo, si dos vectores son representados, respectivamente, por, digamos, $3 + 2i$ y $2 + 4i$, entonces la suma de los números complejos, a saber, $5 + 6i$, representa la suma de los vectores añadidos por medio de la ley del paralelogramo. Lo que los números complejos hacen para los vectores en el plano es proporcionar un álgebra para representar los vectores y las operaciones con vectores. No es necesario realizar las operaciones geométricamente, pero es posible trabajar con ellas algebraicamente, un poco como la ecuación de una curva se emplea para representar (y trabajar con) curvas.

Este empleo de los números complejos para representar vectores en el plano se hizo bien conocido hacia 1830. Sin embargo, la utilidad de los números complejos es limitada. Si varias fuerzas actúan sobre un cuerpo, no están necesariamente en el mismo plano. Para tratarlas algebraicamente es necesario un análogo tridimensional de los números complejos. Se pueden usar las coordenadas cartesianas (x,y,z) de un punto para representar un vector desde el origen al punto, pero no había operaciones con las tripletas de números para representar las operaciones con los vectores. Estas operaciones, como en el caso de los números complejos, deberían incluir la adición, substracción, multiplicación y división y además obedecer a las leyes usuales asociativa, conmutativa y distributiva de tal forma que las operaciones algebraicas se aplicasen con libertad y eficacia. Los ma-

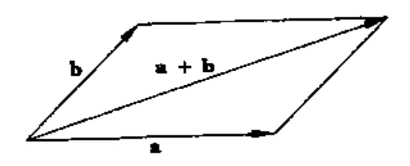


FIGURA 32.1

temáticos iniciaron una búsqueda de lo que fue llamado el número complejo tridimensional y su álgebra.

Wessel, Gauss, Servois, Möbius y otros trabajaron en este problema. Gauss [7] escribió una nota breve e inédita fechada en 1819 sobre mutaciones del espacio. Pensó en los números complejos como desplazamientos; $a + bi$ era un desplazamiento de a unidades a lo largo de una dirección fija, seguido por un desplazamiento de b unidades en una dirección perpendicular. De aquí intentó construir un álgebra de números de tres componentes, en la cual la tercera componente representaría un desplazamiento en una dirección perpendicular al plano de $a + bi$. Llegó a una álgebra no conmutativa, pero no se trataba del álgebra requerida por los físicos. Más aún, ya que no lo publicó, este trabajo tuvo poca influencia.

La creación de un análogo espacial útil de los números complejos se le debe a William R. Hamilton (1805-1865). Después de Newton, Hamilton es el más grande de los matemáticos ingleses y, como a Newton, se le considera aún más grande como físico que como matemático. A la edad de cinco años, Hamilton podía leer latín, griego y hebreo. A los ocho añadió italiano y francés; a los diez podía leer árabe y sánscrito, y a los catorce, persa. Un encuentro con un brillante calculador le inspiró estudiar matemáticas. Entró en el Trinity College, en Dublin, en 1823, y allí destacó como un alumno brillante. En 1822, a la edad de diecisiete años, preparó un artículo sobre caústicas, leído ante la Real Academia Irlandesa dos años después, pero no se publicó, porque se le pidió que trabajara su texto y lo ampliara. En 1827 presentó a la Academia una versión revisada, que tituló «A theory of systems of rays» («Una teoría de sistemas de rayos»), donde hizo una ciencia de la óptica geométrica. Exponía lo que se denominan las funciones características de la óptica. El artículo fue publicado en 1828 en las *Transactions of the Royal Irish Academy* [8].

En 1827, mientras seguía estudiando la licenciatura, recibe el nombramiento de profesor de astronomía del Trinity College, con el título de Astrónomo Real de Irlanda. Sus funciones como profesor consistían en impartir cátedra sobre ciencia y administrar el observatorio astronómico. No realizó un gran trabajo en esto último, pero fue un buen maestro.

---

[7] *Werke*, 8, 357-362.

[8] *Trans. Royal Irish Academy*, 15, 1828, 69-174 = *Math. Papers*, 1, 1-106.

En 1830 y 1832 publicó tres suplementos a «Una teoría de sistema de rayos». En el tercer ensayo [9] predijo que un rayo de luz propagándose en direcciones especiales en un cristal biaxial haría surgir un cono de rayos refractados. Este fenómeno fue confirmado experimentalmente por Humphrey Lloyd, su amigo y colega. Sus ideas en óptica las trasladó a la dinámica y en este campo escribió dos artículos muy famosos (cap. 30), donde aplica el concepto de función característica que había desarrollado en óptica. También aportó un sistema de integrales completas y rigurosas para las ecuaciones diferenciales del movimiento de un sistema de cuerpos. Su principal obra matemática fue sobre los cuaterniones, que discutiremos en breve. La versión final de este trabajo la presentó en sus *Lectures on Quaternions* (*Lecciones sobre Cuaterniones,* 1853) y en dos volúmenes publicados póstumamente, *Elements of Quaternions* (*Elementos de Cuaterniones,* 1866).

Hamilton usaba las analogías con destreza para razonar pasando de lo conocido a lo desconocido. A pesar de que poseía una fina intuición, no tuvo grandes ideas, aunque trabajó dura y largamente en problemas especiales para ver qué aspectos podían ser generalizados. Gozaba de paciencia y era sistemático al analizar muchos ejemplos específicos, y estaba dispuesto a efectuar tremendos cálculos para verificar o demostrar un punto. Sin embargo, en sus publicaciones únicamente se observan los resultados generales pulidos y compactos.

Era profundamente religioso y este interés fue el más importante para él. Luego, en orden de importancia, colocaba la metafísica, matemáticas, poesía, física y literatura general. Escribió poesía, pensaba que las ideas geométricas creadas en su tiempo, el uso de elementos infinitos y elementos imaginarios en los trabajos de Poncelet y Chasles (cap. 35), estaban estrechamente relacionados con la poesía. Aunque era un hombre modesto, admitía y aun enfatizaba que el amor a la fama mueve y motiva a los grandes matemáticos.

La aclaración por Hamilton de la noción de número complejo le permitió pensar con mayor claridad acerca del problema de presentar una analogía tridimensional para representar los vectores en el espacio. Pero el efecto inmediato fue frustrar sus esfuerzos. Todos los números conocidos por los matemáticos de su tiempo poseían la propiedad conmutativa de la multiplicación, y era natural para Ha-

[9] *Trans. Royal Irish Academy,* 17, 1837, 1-144 = *Math. Papers,* 1, 164-293.

milton creer que los números tridimensionales o de tres componentes que él buscaba debían poseer esta misma propiedad al igual que las otras propiedades que poseían los números reales y complejos. Después de algunos años de esfuerzo, Hamilton se vio obligado a hacer dos compromisos. El primero era que sus nuevos números contenían cuatro componentes, y el segundo, que resultaba necesario sacrificar la ley conmutativa de la multiplicación. Ambas características fueron revolucionarias para el álgebra. Llamó a los nuevos números cuaterniones (o cuaternios).

Retrospectivamente, es fácil observar sobre bases geométricas que los nuevos «números» tenían que contener cuatro componentes. El nuevo número, considerado como un operador, se suponía que haría rotar un vector dado alrededor de un eje dado en el espacio, y estirar o contraer el vector. Para estos propósitos, son necesarios dos parámetros (ángulos) a fin de fijar el eje de rotación; un parámetro ha de especificar el ángulo de rotación y el cuarto el estiramiento o contracción del vector dado.

El propio Hamilton describió su descubrimiento de los cuaterniones [10]:

> Mañana será el decimoquinto cumpleaños de los cuaterniones. Ellos surgieron a la vida, o a la luz, ya crecidos, el 16 de octubre de 1843, cuando me encontraba caminando con la Sra. Hamilton hacia Dublín, y llegamos al puente de Broughman. Es decir, entonces y ahí cerré el circuito galvánico del pensamiento y las chispas que cayeron fueron las ecuaciones fundamentales entre *I*, *J*, *K*; *exactamente como* las he usado desde entonces. Saqué, en ese momento, una libreta de bolsillo, que todavía existe, e hice una anotación, sobre la cual, *en ese mismo preciso momento*, sentí que posiblemente sería valioso el extender mi labor por al menos los diez (o podían ser quince) años por venir. Pero entonces era justo decir que esto sucedía porque sentí en ese momento que un *problema* había sido resuelto, un deseo intelectual *aliviado*, deseo que me había *perseguido* por lo menos los *quince años* anteriores.

Anunció la invención de los cuaterniones en 1843, durante una reunión de la Real Academia Irlandesa, y después dedicó el resto de su vida a desarrollar la teoría, y escribió muchos ensayos en torno a ellos.

---

[10] *North British Review*, 14, 1858, 57.

3. La naturaleza de los cuaterniones

Un cuaternión es un número de la forma

$$3 + 2\mathbf{i} + 6\mathbf{j} + 7\mathbf{k} \tag{2}$$

donde las *i*, *j* y *k* tienen un papel semejante al que juega *i* en los números complejos. La parte real, 3 arriba, es llamada la parte escalar del cuaternión, y el resto, la parte vectorial. Los tres coeficientes de la parte vectorial son coordenadas rectangulares cartesianas de un punto *P*, mientras que *i*, *j* y *k* son llamadas las unidades cualitativas que geométricamente están dirigidas a lo largo de los tres ejes. El criterio de igualdad de dos cuaterniones es que sus partes escalares deben ser iguales y que los coeficientes de sus unidades, *i*, *j* y *k* han de ser respectivamente iguales. Dos cuaterniones se suman sumando sus partes escalares y sumando los coeficientes de cada una de las unidades *i*, *j* y *k* para formar los nuevos coeficientes de estas unidades. La suma de dos cuaternios es, por lo tanto, otro cuaternio.

Todas las reglas algebraicas familiares de multiplicación son supuestamente válidas al operar con cuaterniones, excepto que al formar productos de estas unidades *i*, *j* y *k*, las siguientes reglas, que abandonan la ley de conmutatividad, se cumplen

$$\mathbf{jk} = \mathbf{i}, \quad \mathbf{kj} = -\mathbf{i}, \quad \mathbf{ki} = \mathbf{j}, \quad \mathbf{ik} = -\mathbf{j}, \quad \mathbf{ij} = \mathbf{k}, \quad \mathbf{ji} = -\mathbf{k}$$
$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1. \tag{3}$$

Así, si

$$\mathbf{p} = 3 + 2\mathbf{i} + 6\mathbf{j} + 7\mathbf{k} \quad \text{y} \quad \mathbf{q} = 4 + 6\mathbf{i} + 8\mathbf{j} + 9\mathbf{k}$$

entonces

$$\mathbf{pq} = (3+2\mathbf{i}+6\mathbf{j}+7\mathbf{k})(4+6\mathbf{i}+8\mathbf{j}+9\mathbf{k}) = -111+24\mathbf{i}+72\mathbf{j}+75\mathbf{k}.$$

mientras que

$$\mathbf{qp} = (4 + 6\mathbf{i} + 8\mathbf{j} + 9\mathbf{k})(3 + 2\mathbf{i} + 6\mathbf{j} + 7\mathbf{k}) = -111 + 28\mathbf{i} + 24\mathbf{j} + 75\mathbf{k}.$$

Hamilton demostró que la multiplicación es asociativa. Este es el primer uso de ese término [11].

La división de un cuaternión por otro puede ser efectuada también, pero el hecho de que la multiplicación no es conmutativa implica que dividir un cuaternión p por un cuaternión q puede significar encontrar r tal que $p = qr$ o tal que $p = rq$. El cociente r no necesita ser el mismo en los dos casos. El problema de la división se maneja mejor al introducir q^{-1} o $1/q$. Si $q = a + bi + cj + dk$ se define q' como $a - bi - cj - dk$ y $N(q)$, llamado la norma de q, como $a^2 + b^2 + c^2 + d^2$. Entonces $N(q) = qq' = q'q$. Por definición $q^{-1} = q'/N(q)$ y q^{-1} existe si $N(q) \neq 0$. También $qq^{-1} = 1$ y $q^{-1}q = 1$. Ahora, para encontrar la r tal que $p = qr$ tenemos $q^{-1}p = q^{-1}qr$ o bien $r = q^{-1}p$. Para encontrar la r tal que $p = rq$ tenemos $pq^{-1} = rqq^{-1}$ o $r = pq^{-1}$.

Que los cuaterniones pueden ser usados para girar y estirar o contraer un vector dado en otro vector dado se demuestra fácilmente. Sólo se debe mostrar que es posible determinar las a, b, c y d tales que

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}).$$

Multiplicando el lado izquierdo como cuaterniones e igualando los coeficientes correspondientes de los lados derecho e izquierdo obtenemos cuatro equaciones en las incógnitas a, b, c y d. Estas cuatro ecuaciones son suficientes para determinar las incógnitas.

Hamilton introdujo también un operador diferencial importante. El símbolo ∇, que es una Δ invertida —que Hamilton llamó «nabla» porque se asemeja a un antiguo instrumento musical hebreo de ese nombre— se emplea para el operador

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}. \tag{4}$$

Cuando se le aplica a una función escalar de punto $u(x, y, z)$ da el vector

$$\nabla u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}. \tag{5}$$

[11] *Proc. Royal Irish Academy*, 2, 1844, 424-434 = *Math. Papers*, 3, 111-116.

Este vector, que varía de un punto a otro del espacio, es ahora llamado el gradiente de u. Representa en magnitud y dirección la máxima rapidez (en espacio) de incremento de u.

También, haciendo $v = v_1i + v_2j + v_3k$ una función de punto continua, donde las v_1, v_2 y v_3, son funciones de x, y y z, Hamilton introdujo

$$\nabla \mathbf{v} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k})$$

$$= -\left(\frac{\Sigma v_1}{\partial x}+\frac{\Sigma v_2}{\partial y}+\frac{\Sigma v_3}{\partial k}\right)+\left(\frac{\Sigma v_3}{\partial y}-\frac{\Sigma v_2}{\partial z}\right)\mathbf{i}+\left(\frac{\Sigma v_1}{\partial z}-\frac{\Sigma v_2}{\partial z}\right)\mathbf{j} \qquad (6)$$

$$= +\left(\frac{\Sigma v_2}{\partial x}-\frac{\Sigma v_1}{\partial y}\right)\mathbf{k}.$$

Así, el resultado de operar con ∇ sobre una función vectorial de punto v da un cuartenión; la parte escalar de este cuaternión (excepto por el signo negativo) es lo que llamamos la divergencia de v y la parte vectorial es llamada ahora rot v.

El entusiasmo de Hamilton por los cuaterniones no tuvo límites. Creía que su invención era tan importante como el cálculo y que sería el instrumento clave en la física matemática. El mismo llevó a cabo algunas aplicaciones a la geometría, óptica y mecánica. Sus ideas fueron apoyadas entusiastamente por su amigo Peter Guthrie Tait (1831-1901), profesor de matemáticas en el Queen's College y posteriormente profesor de historia natural en la universidad de Edinburgo. En muchos artículos, Tait motivó a los físicos a adoptar los cuaterniones como herramienta básica. Incluso se vio envuelto en largas discusiones con Cayley, quien tomó un punto de vista muy tibio en cuanto a la utilidad de los cuaterniones. Pero los físicos ignoraron los cuaterniones y continuaron trabajando con las coordenadas cartesianas convencionales. Sin embargo, como veremos más adelante, el trabajo de Hamilton condujo indirectamente a un álgebra y análisis de vectores que los físicos adoptaron ansiosamente.

Los cuaterniones de Hamilton demostraron ser de importancia inconmensurable para el álgebra. Una vez que los matemáticos se dieron cuenta de que un sistema de números útil y con sentido podía ser construido y además carecer de la propiedad conmutativa de los

números reales y complejos, se sintieron más libres para considerar creaciones que se alejaban aún más de las propiedades usuales de los números reales y complejos. Esta conciencia era necesaria antes de que el álgebra y análisis vectorial pudieran ser desarrollados, porque los vectores violan aún más leyes ordinarias del álgebra de que lo que lo hacen los cuaterniones (sec. 5). Aún más general, el trabajo de Hamilton condujo a la teoría de álgebras lineales asociativas (sec. 6). El propio Hamilton inició un trabajo sobre los hipernúmeros que contenían n componentes o n-uplas [12], pero fue su trabajo sobre los cuaterniones lo que estimuló los nuevos trabajos sobre las álgebras lineales.

4. El cálculo de la extensión de Grassmann

Mientras que Hamilton desarrollaba sus cuaterniones, otro matemático, Hermann Gunther Grassmann (1809-1877), aunque no mostrara talento matemático en su juventud y careciera de educación universitaria en matemáticas, pero que más tarde se convirtió en maestro del *gymnasium* (instituto) en Stettin, Alemania, así como en una autoridad en sánscrito, estaba desarrollando una generalización más audaz de los números complejos. Grassmann tuvo sus ideas antes que Hamilton, pero no las publicó hasta 1844, un año después de que Hamilton anunciara su descubrimiento de los cuaterniones. En ese año publicó su *Die lineale Ausdehnungoslehre (El cálculo de la extensión)*. Debido a que cubrió las ideas con doctrinas místicas y a que su exposición fue abstracta, los matemáticos de mentalidad más práctica y los físicos encontraron su trabajo vago y poco legible; como consecuencia, el trabajo, aunque era altamente original, permaneció casi en el anonimato por varios años. Grassmann publicó una edición revisada: *Die Ausdehnungslehre (La extensión)*, en 1862, donde simplificó y amplió el trabajo original, pero su estilo y carencia de claridad repelieron de nuevo a los lectores.

A pesar de que la exposición de Grassmann estaba casi inextricablemente ligada con ideas geométricas —de hecho, él estaba interesado en la geometría— abstraeremos las nociones algebraicas que han probado ser de incalculable valor. Su noción básica, que llamó

---

[12] *Trans. Royal Irish Academy*, 21, 1848, 199-296 = *Math. Papers*, III, 159-226

cantidad extensiva *(extensive Grösse)* es un tipo de hipernúmero con n componentes. Para estudiar sus ideas discutiremos el caso $n = 3$.

Consideremos dos hipernúmeros

$$\alpha = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_2 e_2 + \alpha_3 e_3 \quad \text{y} \quad \beta = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3,$$

donde las α_i y β_i son números reales y donde e_1, e_2 y e_3 son unidades cualitativas o primarias representadas geométricamente por segmentos de línea dirigidos (de una unidad de longitud) trazados desde un origen común determinando un sistema de ejes ortogonal orientado a derechas. Las $\alpha_i e_i$ son múltiplos de las unidades primarias y están representadas geométricamente por longitudes α_i a lo largo de los ejes respectivos, mientras que α está representada por un segmento de línea dirigido en el espacio cuyas proyecciones sobre los ejes son las longitudes α_i. Lo mismo es cierto para las β_i y β. Grassmann llamó a los segmentos de recta dirigidos o vectores-línea, *Strecke*.

La adición y sustracción de estos hipernúmeros está definida mediante

$$\alpha \pm \beta = (\alpha_1 \pm \beta_1)e_1 + (\alpha_2 \pm \beta_2)e_2 + (\alpha_3 \pm \beta_3)e_3. \tag{7}$$

Grassmann introdujo dos clases de multiplicaciones, el producto interno y el producto externo. Para el producto interno postuló que

$$e_i|e_i = 1, \; e_i|e_j = 0 \qquad \text{para } i \neq j. \tag{8}$$

Para el producto externo

$$[e_i e_j] = -[e_j e_i], \; [e_i e_i] = 0. \tag{9}$$

Estos corchetes se llaman unidades de segundo orden y no son reducidos por Grassmann (mientras que Hamilton sí lo hace) a unidades de primer orden, esto es, a las e_i, pero las trata como si fueran equivalentes a las unidades de primer orden, con $[e_1 e_2] = e_3$, y así sucesivamente.

A partir de estas definiciones se sigue que el producto interno $\alpha|\beta$ de α y β está dado por

$$\alpha|\beta = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 \;\text{ y }\; \alpha|\beta = \beta|\alpha.$$

El valor numérico o magnitud a de un hipernúmero α está definido como $\sqrt{\alpha|\alpha} = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$. Así, la magnitud de α es igual numéricamente a la longitud del vector-línea que lo representa geométricamente. Si θ denota el ángulo entre los vectores-línea α y β, entonces

$$\alpha|\beta = ab\left(\frac{\alpha_1\beta_1}{ab} + \frac{\alpha_2\beta_2}{ab} + \frac{\alpha_3\beta_3}{ab}\right) = ab \cos \theta.$$

Con la ayuda de la regla del producto externo (9), el producto externo P de los hipernúmeros α y β puede ser expresado como sigue:

$$P = [\alpha\beta] = (\alpha_2\beta_3 - \alpha_3\beta_2)[e_2e_3] + (\alpha_3\beta_1 - \alpha_1\beta_3)[e_3e_1] + (\alpha_1\beta_2 - \alpha_2\beta_1)[e_1e_2]. \quad (10)$$

Este producto es un hipernúmero de segundo orden y está expresado en términos de unidades independientes de segundo orden. Su magnitud $|P|$ se obtiene mediante una definición del producto interno de dos hipernúmeros de segundo orden y es

$$|P| = \sqrt{P|P} = \{(\alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_3\beta_1 - \alpha_1\beta_3)^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)^2\}^{1/2}$$

$$= ab\left\{1 - \left(\frac{\alpha_1\beta_1}{ab} + \frac{\alpha_2\beta_2}{ab} + \frac{\alpha_3\beta_3}{ab}\right)^2\right\}^{1/2} \quad (11)$$

$$= ab \text{ sen } \theta.$$

Por tanto la magnitud $|P|$ del producto interno $[\alpha\beta]$ está representada geométricamente por el área del paralelogramo construido sobre los vectores-líneas que son las representaciones geométricas de α y β. Este área, junto con una unidad de vector-línea normal a él, cuya dirección es seleccionada de tal forma que si α es rotado hacia β alrededor de la normal, entonces la normal señalará la dirección de un tornillo rotando hacia la derecha de α a β, es ahora llamado un área vectorial. El término de Grassmann fue *Plangrösse*.

El producto interno de Grassmann de dos hipernúmeros primarios para tres dimensiones es equivalente a (con signo menos) la parte escalar del producto de cuaterniones de Hamilton de dos vectores; y, de nuevo en el caso tridimensional, el producto externo de Grassmann, si reemplazamos $[e_2e_3]$ por e_1, y así con los demás, es

precisamente el producto de cuaterniones de Hamilton de dos vectores. Sin embargo, en la teoría de los cuaterniones, el vector es una parte subsidiaria del cuaternión mientras que en el álgebra de Grassmann el vector aparece como la cantidad básica.

Otro producto fue formado por Grassmann al tomar el producto (interno) escalar de un hipernúmero γ con el producto (exterior) vectorial de dos hipernúmeros α y β. Este producto Q es para el caso tridimensional

$$Q=[\alpha\beta]\gamma=(\alpha_2\beta_3-\alpha_3\beta_2)\gamma_1+(\alpha_3\beta_1-\alpha_1\beta_3)\gamma_2+(\alpha_1\beta_2-\alpha_2\beta_1)\gamma_3.$$

En forma de determinante

$$Q=\begin{vmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{vmatrix}. \qquad (12)$$

Consecuentemente, Q puede ser interpretado geométricamente como el volumen de un paralelepípedo construido sobre los vectores-línea que representan α, β y γ. El volumen puede ser positivo o negativo.

Grassmann consideró (para hipernúmeros de n componentes) no solamente los dos tipos de productos descritos con anterioridad, sino también productos de orden superior. En un ensayo de 1855 [13], proporcionó dieciséis tipos diferentes de productos para hipernúmeros. Además, dio el significado geométrico de los productos e hizo aplicaciones a la mecánica, magnetismo y cristalografía.

Parece como si el tratamiento de Grassmann de los hipernúmeros de n partes fuese innecesariamente general, ya que hasta ahí, al menos las partes útiles de los hipernúmeros, contenían como mucho cuatro partes. Aun así, el pensamiento de Grassmann llevó a los matemáticos a la teoría de los tensores (cap. 48), que, como veremos, son hipernúmeros. Otras ideas geométricas y la noción de invariancia tendrían que ser conocidas por los matemáticos antes que llegara el día de los tensores. A pesar de que la concepción de los hipernúmeros condujo a varias generalizaciones, no hubo ningún desarrollo del análisis (*e.g.*, cálculo) de los hipernúmeros n-dimensionales de

---

[13] *Jour. für Math.*, 49, 1855, 10-20 y 123-141 = *Ges. Math. und Phys. Werke*, 2, Parte I, 199-217.

Grassmann. La razón es simplemente que nunca se encontraron aplicaciones para tal tipo de análisis. Como veremos, existe un extenso análisis para tensores, pero éstos tienen su origen en la geometría riemanniana.

5. De los cuaterniones a los vectores

La obra de Grassmann permaneció ignorada por un tiempo, mientras que, como hemos anotado, los cuaterniones recibieron gran atención casi de inmediato. Sin embargo, no eran lo que los físicos realmente necesitaban. Ellos buscaban un concepto que no estuviera disociado, sino, al contrario, más relacionado con las coordenadas cartesianas de lo que lo estaban los cuaterniones. El primer paso en la dirección de tal concepto lo dio James Clerk Maxwell (1831-1879), fundador de la teoría electromagnética, uno de los más grandes físicos matemáticos y profesor de física en la universidad de Cambridge.

Maxwell conoció el trabajo de Hamilton y aunque supo del trabajo de Grassmann, no lo había visto. Singularizó las partes escalar y vectorial de los cuaterniones de Hamilton y puso énfasis sobre estas nociones separadas [14]. Sin embargo, en su famoso *A Treatise on Electricity and Magnetism* (*Tratado de electricidad y magnetismo,* 1873) hizo mayores concesiones a los cuaterniones y habla de las partes escalar y vectorial de un cuaternión, aunque las trata como entidades separadas. Un vector —dice (p. 10)— requiere tres cantidades (componentes) para su especificación que se interpretan como longitudes a lo largo de los tres ejes coordenados. Este concepto de vector es la parte vectorial de los cuaterniones de Hamilton, y así lo expone Maxwell. Hamilton había introducido una función vectorial v de x, y y z con componentes v_1, v_2 y v_3, había aplicado el operador $\nabla = i(\partial/\partial x) + j(\partial/\partial y) + k(\partial/\partial z)$, y obtenido el resultado (6). Por tanto ∇_v es un cuaternio. Pero Maxwell separó la parte escalar de la parte vectorial e indicó éstas por $S\nabla v$ (la parte escalar de v) y $V\nabla v$ (la parte vectorial de v). Llamó a $S\nabla v$ la convergencia de v, ya que esta expresión había aparecido muchas veces en la mecánica de fluidos y cuando v es la velocidad tiene el significado

---

[14] *Proc. London Math. Soc.*, 3, 1871, 224-232 = *The Scientific Papers*, vol. 2, 257-266.

de flujo o cantidad neta por unidad de volumen por unidad de tiempo que fluye a través de una pequeña área alrededor de un punto. Y llamó a $V\nabla v$ el rotacional de v, porque la expresión estaba también acuñada en la mecánica de fluidos como dos veces la velocidad de rotación del fluido en un punto. Clifford, más tarde, llamo a $-S\nabla v$ divergencia.

Maxwell señala entonces que el operador ∇ repetido proporciona

$$\nabla^2 = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2},$$

lo que él denomina operador de Laplace. Esto le permite actuar sobre una función escalar para obtener un escalar y sobre una función vectorial para obtener un vector [15].

Maxwell notó en su artículo de 1871 que el rotacional de un gradiente de una función escalar y la divergencia de rotacional de una función vectorial son siempre cero. También establece que el rotacional del rotacional de una función vectorial v es el gradiente de la divergencia de v menos el laplaciano de v. (Esto es cierto únicamente en coordenadas rectangulares.)

Maxwell usó comúnmente los cuaterniones como la entidad matemática básica o al menos hizo frecuentes referencias a los cuateriones, tal vez para ayudar a sus lectores. Sin embargo, su trabajo aclaró que los vectores eran la verdadera herramienta para el pensamiento físico y no sólo un esquema abreviado de escritura, como alguien mantenía. De este modo fue creada en el tiempo de Maxwell una gran parte del análisis vectorial al tratar las partes escalares y vectoriales de los cuaterniones por separado.

La ruptura formal con los cuaterniones y la inauguración de una

---

[15] Ya que en análisis vectorial $\nabla^2 = \nabla . \nabla$, entonces $\nabla^2 q$ significa físicamente la divergencia del gradiente o la divergencia de la máxima razón de cambio en el espacio de q. Sin embargo, el significado físico es más claro a partir del hecho de que la función q que satisface $\nabla^2 q = 0$ minimiza la integral de Dirichlet (véase (34) del cap. 28). La integral es el cuadrado de la magnitud del gradiente tomada sobre algún volumen. De aquí que $\nabla^2 q = 0$ significa que el gradiente es mínimo en cualquier punto o que la distancia de la uniformidad es mínima. Si q no es cero entonces debe haber alguna separación de la uniformidad y habrá una fuerza restauradora. Las diversas ecuaciones de la física matemática, que contienen a $\nabla^2 q$ en un contexto u otro, en efecto, establecen que la naturaleza siempre actúa para restaurar la uniformidad.

La notación Δ para ∇^2 fue introducida por Robert Murphy en 1833.

nueva materia independiente, el análisis vectorial tridimensional, fue hecho independientemente por Josiah Willard Gibbs y Oliver Heaviside a principio de los 1880. Gibbs (1839-1903), profesor de matemáticas en Yale College, pero principalmente un físico-químico, había impreso (1881-1884) para distribución privada entre sus estudiantes un panfleto sobre los *Elements of vector analysis (Elementos del Análisis Vectorial)* [16]. Su punto de vista se expone en una nota introductoria:

> Los principios fundamentales del siguiente análisis son familiares bajo una forma ligeramente diferente a los estudiantes de los cuaterniones. La manera como la materia es desarrollada de alguna manera es diferente de lo que se hace en los tratados sobre cuaterniones, consistiendo simplemente en dar una notación adecuada para aquellas relaciones entre los vectores, o entre vectores y escalares, que parecen más importantes, y que se prestan ellas mismas más fácilmente a las transformaciones analíticas, y para explicar algunas de estas transformaciones. Como precedente de tal desviación del uso de los cuaterniones puede ser citado el *Kinematics (Kinemática)* de Clifford. En este contexto puede ser mencionado el mismo Grassmann, a cuyos sistemas se asemeja el siguiente método en algunos aspectos más cercanamente que el de Hamilton.

Aunque impreso para circulación privada, el panfleto de Gibbs sobre análisis vectorial fue ampliamente conocido. El material se incorporó finalmente en un libro escrito por E. B. Wilson y basado sobre las clases de Gibbs. El libro de Gibbs y Wilson titulado *Vector Analysis (Analisis Vectorial)* apareció en 1901.

Oliver Heaviside (1850-1925) fue en la primera parte de su carrera científica un ingeniero que aplicó su trabajo a telégrafos y teléfonos. Se retiró a la vida campestre en 1874, y se dedicó a escribir, principalmente sobre electricidad y magnetismo. Heaviside estudió los cuaterniones en los *Elementos* de Hamilton, pero había sido repelido por tantos teoremas particulares. Pensaba que los cuaterniones eran demasiado difíciles de aprender para ingenieros ocupados y construyó su propio análisis vectorial, que para él era únicamente una forma abreviada de las coordenadas cartesianas ordinarias. En artículos escritos durante los 1880 en la revista *Electrician*, usó su análisis vectorial libremente. Más adelante, en su trabajo en tres volúmenes *Electromagnetic Theory* (*Teoría Electromagnética*, 1893,

---

[16] *The Scientific Papers*, 2, 17-90.

1899 y 1912) aportó gran cantidad de álgebra vectorial en el vol. I. El tercer capítulo, de alrededor de 175 páginas, está dedicado a los métodos vectoriales. Su desarrollo de la materia está esencialmente en armonía con el de Gibbs, aunque no le gustaba la notación de Gibbs y adoptó la suya propia basada en la notación cuaterniónica de Tait.

Según la formulación de Gibbs y Heaviside, un vector no es más que la parte vectorial de un cuaternión, pero considerado independientemente de cualesquiera cuaterniones. Así un vector v es

$$v = ai + bj + ck$$

donde i, j y k son vectores unitarios a lo largo de los ejes x, y y z, respectivamente. Los coeficientes a, b y c son números reales y se les llama componentes. Dos vectores son iguales si las componentes respectivas son iguales y la suma de dos vectores es el vector que tiene como componentes la suma de las componentes respectivas de los sumandos.

Se introdujeron dos tipos de multiplicación, ambas útiles físicamente. El primer tipo, conocido como multiplicación escalar, se define así: multiplicamos v y $v' = a'i + b'j + c'k$ como polinomios ordinarios y, usando el punto como un símbolo de multiplicación, fijamos

$$\begin{aligned} \mathbf{i}\cdot\mathbf{i} &= \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1 \\ \mathbf{i}\cdot\mathbf{j} &= \mathbf{j}\cdot\mathbf{i} = \mathbf{i}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{j} = 0. \end{aligned} \tag{13a}$$

Así $v.v' = aa' + bb' + cc'$. Este producto ya no es un vector, sino un número real, o escalar, y es llamado producto escalar. Posee una característica algebraica nueva, ya que el producto de dos números reales o complejos o cuaterniones es siempre un número de la misma clase con que empezamos. Otra propiedad sorprendente del producto escalar es que puede ser cero cuando ninguno de los factores es cero. Por ejemplo, el producto de los vectores $v = 3i$ y $v' = 6j + 7k$ es cero.

El producto escalar de dos vectores es novedoso algebraicamente en otro aspecto —no permite un proceso inverso—. Esto es, no siempre podemos encontrar un vector o un escalar q tal que $v/v' = q$. Así, si q fuera un vector, $q.v'$ sería un escalar y no sería igual al vector v. Por otra parte, si q fuera un escalar, entonces, a pesar de

que qv' está definido como $qa'i + qb'j + qc'k$, sería raro que $qa' = a$, $qb' = b$ y $qc' = c$ donde a, b y c son los coeficientes de v. A pesar de la ausencia de un cociente, el producto escalar es útil.

La interpretación física del producto escalar es fácilmente expresable como lo siguiente: si v' es una fuerza (Fig. 32.2) cuya dirección y magnitud están representadas por el segmento de recta de O a P', entonces la efectividad de la fuerza al empujar un objeto en O, digamos, en la dirección de OP, donde OP representa a v, es la proyección de OP' sobre OP o bien OP' cos ϕ donde ϕ es el ángulo entre OP' y OP. Si OP es de una unidad de longitud, la proyección de OP' es precisamente el valor del producto $v.v'$.

La segunda clase de producto de vectores, llamado producto vectorial, se define como sigue: multiplicamos de nuevo v y v' como polinomios pero esta vez hacemos

$$\begin{gathered}\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0 \\ \mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{i} = -\mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \\ \mathbf{k} \times \mathbf{i} = \mathbf{j}, \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}.\end{gathered} \tag{13b}$$

Entonces el producto, que se indica por $v \times v'$, es

$$\mathbf{v} \times \mathbf{v}' = (bc' - b'c)\mathbf{i} + (ca' - ac')\mathbf{j} + (ab' - b'a)\mathbf{k}.$$

El producto vectorial de dos vectores es un vector. Su dirección puede demostrarse que es perpendicular a las de v y v' y señala la

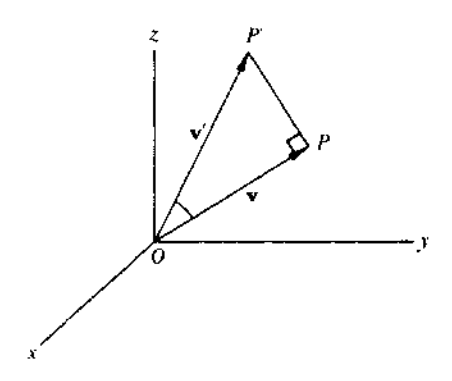


FIGURA 32.2

dirección que un tornillo a derechas se mueve cuando es girado de v a v' siguiendo el menor ángulo. El producto vectorial de dos vectores paralelos es cero, aunque ningún factor lo es. Más aún, este producto, como el producto de cuaterniones, no es conmutativo. Además, no es ni siquiera asociativo. Por ejemplo, $i \times j \times j$ puede significar $(i \times j) \times j = k \times j = -i$ ó $i \times (j \times j) = i \times 0 = 0$.

No existe el inverso de la multiplicación vectorial. Para que el cociente de v por v' sea un vector q tendríamos que tener

$$v = v' \times q.$$

Y esto requiere, cualquiera que sea q, que v' sea perpendicular a v, lo que podría no ser el caso para empezar. Si q fuera un escalar, sería accidental que $qa'i + qb'j + qc'k$ fuera igual a v.

El producto vectorial, como el producto escalar, es sugerido por situaciones físicas. Sean OP y PP' en la Figura 32.3 las longitudes y direcciones de v y v'. Si v' es una fuerza cuya magnitud y dirección son las de PP', la medida del momento de la fuerza ejercida por v' alrededor de O es la longitud del vector $v \times v'$ y en general es tomado como si tuviera la dirección de $v \times v'$.

El álgebra de vectores es extendida a vectores variables y a un cálculo de vectores. Por ejemplo, el vector variable $v(t) = a(t)i + b(t)j + c(t)k$, donde $a(t)$, $b(t)$ y $c(t)$ son funciones de t, es una función vectorial. Si los diversos vectores que se obtiene para valores distintos de t son trazados a partir de O como origen (Fig. 32.4), los extremos de estos vectores describen una curva. De aquí que la función vectorial de una variable escalar t tenga un papel análogo al de la función ordinaria

$$y = x^2 + 7,$$

digamos, y el cálculo de vectores ha sido desarrollado para estas funciones vectoriales, así como hay uno para funciones ordinarias.

Los conceptos de gradiente de u

$$\nabla u = \frac{\partial u}{\partial x}\,\mathbf{i} + \frac{\partial u}{\partial y}\,\mathbf{j} + \frac{\partial u}{\partial z}\,\mathbf{k}, \qquad (14)$$

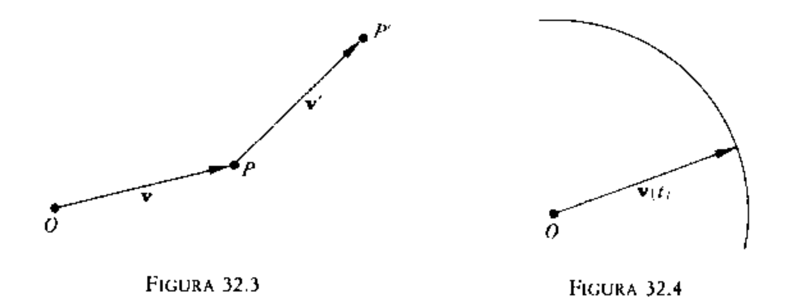


FIGURA 32.3

FIGURA 32.4

donde u es una función escalar de x, y y z, la divergencia de una función vectorial v,

$$\nabla \cdot \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}, \tag{15}$$

donde v_1, v_2 y v_3 son las componentes de v, y el rotacional de v,

$$\nabla \times \mathbf{v} = \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}, \tag{16}$$

fueron abstraídos de los cuaterniones.

Muchos teoremas básicos del análisis pueden ser expresados en forma vectorial. Así, al resolver las ecuaciones diferenciales parciales del calor, Ostrogradsky [17] hizo uso de la siguiente conversión de la integral de volumen a la integral de superficie:

$$\iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx dy dz = \iint_S (P\cos\lambda + Q\cos\mu + R\cos v) dS.$$

Aquí P, Q y R son funciones de x, y y z y son componentes de un vector, λ, μ, y v son los cosenos direccionales de la normal a la superficie S que limita el volumen V sobre el que la integral de la izquierda es calculada. Este teorema, conocido como el teorema de divergencia (también como teorema de Gauss y teorema de Ostrogradsky) puede ser expresado en forma vectorial así: si F es un vec-

[17] *Mem. Acad. Sci. St. Peters.* (6), 1, 1831, 39-53.

tor cuyas componentes son P, Q y R y n es la normal a S entonces

$$\iiint \nabla \cdot \mathbf{F}\, dv = \iint \mathbf{F} \cdot \mathbf{n}\, dS. \tag{17}$$

Del mismo modo el teorema de Stokes, que fue primero establecido por él como una pregunta en un examen para el premio Smith en Cambridge en 1854 [18], establece en forma escalar que

$$\iint_S \left\{\lambda\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \mu\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + v\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right\} dS$$
$$= \int_C \left(P\,\frac{\partial x}{\partial s} + Q\,\frac{\partial y}{\partial s} + R\,\frac{\partial z}{\partial s}\right) ds,$$

donde S es cualquier porción de una superficie, C es la curva limitando S, y $x(x)$, $y(s)$ y $z(s)$ son la representación paramétrica de C. En forma vectorial el teorema de Stokes se enuncia

$$\iint \text{rot}\, \mathbf{F} \cdot \mathbf{n}\, dS = \int \mathbf{F} \cdot \frac{d\mathbf{r}}{ds}\, ds, \tag{18}$$

donde $r(s)$ es el vector cuyas componentes son $x(s)$, $y(s)$ y $z(s)$.

Cuando Maxwell escribió las expresiones y ecuaciones del electromagnetismo, especialmente las ecuaciones que ahora llevan su nombre, las escribió en general separando las componentes de los vectores involucrados en grad u, div v, y rot v. Sin embargo, Heaviside escribió las ecuaciones de Maxwell en forma vectorial (cap. 28 (52)).

Es cierto que los cálculos con vectores y funciones vectoriales se hacen frecuentemente apoyándose en las componentes cartesianas, pero es altamente importante pensar también en términos de una entidad singular, el vector, y en términos de gradiente, divergencia y rotacional. Estos tienen un significado físico directo, por no decir nada del hecho de que los pasos técnicos complicados pueden realizarse directamente con los vectores, como cuando nos reemplaza $\nabla.(u(x,y,z) \times v\,(x,y,z))$ por su equivalente $v.\nabla x u - u.\nabla x v$. También se habían dado definiciones integrales del gradiente, divergencia y

[18] El teorema fue establecido por lord Kelvin en una carta a Stokes de julio de 1850.

rotacional, haciendo estos conceptos independientes de cualquier definición de coordenadas. Así, en lugar de (14) tenemos, por ejemplo,

$$\text{grad } u = \lim_{\Delta t \to 0} \frac{1}{\Delta \tau} \int_S u \, n \, dS$$

donde S es la frontera de un elemento de volumen Δr y n es la normal al elemento de superficie dS de S.

Poco después que el análisis vectorial fuera creado, surgieron muchas controversias entre los que proponían los cuaterniones y los que postulaban los vectores, a fin de decidir qué resultaba más útil. Los cuaternionistas fueron fanáticos en cuanto al valor de los cuaterniones, pero los que proponían el análisis vectorial se mostraron igualmente partidistas. En un lado estaban alineados los grandes defensores de los cuaterniones, tales como Tait; en el otro, Gibbs y Heaviside. A propósito de la controversia, Heaviside señaló sarcásticamente que para el tratamiento de los cuaterniones, los cuaterniones son el mejor instrumento. Por otro lado, Tait descubrió el álgebra vectorial de Heaviside como «un tipo de monstruo hermafrodita, compuesto por las notaciones de Grassmann y Hamilton». El libro de Gibbs demostró ser de un valor inestimable para promover la causa de los vectores.

El asunto se decidió finalmente en favor de los vectores. Los ingenieros recibieron bien el análisis vectorial de Gibbs y Heaviside, aunque los matemáticos no. Para los inicios del presente siglo, los físicos también estaban bastante convencidos de que el análisis vectorial era lo que ellos necesitaban. Los libros de texto sobre la materia aparecieron pronto en todos los países y son ahora comunes. Finalmente, los matemáticos siguieron e introdujeron los métodos vectoriales en las geometrías analítica y diferencial.

La influencia de la física en estimular la creación de entidades matemáticas tales como los cuaterniones, los hipernúmeros de Grassmann y los vectores debe considerarse importante. Estas creaciones se hicieron parte de las matemáticas. Pero su significado se extiende más allá de la adición de nuevas materias. La introducción de estas distintas cantidades abrió una nueva visión —no solamente hay una álgebra de los números reales y complejos, sino muchas álgebras diversas.

6. Algebras lineales asociativas

Desde el punto de vista puramente algebraico, los cuaterniones eran atractivos, ya que proporcionaban un ejemplo de un álgebra que tenía las propiedades de los números reales y complejos excepto para la conmutatividad de la multiplicación. Durante la segunda mitad del siglo XIX, fueron explorados muchos sistemas de hipernúmeros, en parte para ver qué variedades podían crearse, y aun retener muchas propiedades de los números reales y complejos.

Cayley [19] proporcionó una generalización de ocho unidades de cuaternios reales. Sus unidades fueron *1*, e_1, e_2,...,e_7, con

$$e_i^2 = -1, \qquad e_ie_j = -e_je_i \quad \text{para } i, j = 1, 2, ..., 7 \quad \text{e} \quad i \neq j$$

$$e_1e_2 = e_3, \qquad e_1e_4 = e_5, \qquad e_1e_6 = e_7 \; .$$

$$e_2e_5 = e_7, \qquad e_2e_4 = -e_6, \qquad e_3e_4 = e_7, \qquad e_3e_5 = e_6$$

y las catorce ecuaciones obtenidas a partir de estas últimas siete al permutar cada conjunto de tres subíndices cíclicamente: *e.g.*, $e_2e_3 = e_1$; $e_3e_1 = 3_2$.

Un número general (octoniano) x está definido por

$$x = x_0 + x_1e_1 + ... + x_qe_q.$$

donde las x_i son números reales. La norma de x, $N(x)$, es por definición

$$N(x) = x_0^2 + x_1^2 + ... + x_q^2.$$

La norma del producto iguala el producto de las normas. La ley asociativa de la multiplicación falla en general (como lo hace la ley conmutativa de la multiplicación). Las divisiones a derecha e izquierda, excepto por cero, son siempre posibles y únicas, un hecho éste pasado por alto por Cayley y demostrado por Leonard Eugene Dickson [20]. En artículos posteriores, Cayley proporcionó otras álgebras de hipernúmeros diferiendo en algo de la anterior.

---

[19] *Phil. Mag.* (3), 26, 1845, 210-213 y 30, 1847, 257-258 = *Coll. Math. Papers*, 1, 127 y 301.

[20] *Amer. Math. Soc. Trans.*, 13, 1912, 59-73.

Hamilton en su *Lecciones sobre cuaterniones* [21] introdujo también los bicuaterniones, esto es, cuaterniones con coeficientes complejos. Notó que la ley del producto no se mantiene para los bicuaterniones; esto es, dos bicuaterniones diferentes de cero pueden tener un producto igual a cero.

William Kingdom Clifford (1845-1879), profesor de matemáticas y mecánica del University College de Londres, creó otro tipo de hipernúmeros [22], que también llamó bicuaterniones. Si q y Q son cuaterniones reales, si w es tal que $w^2 = 1$, y si w conmuta con todo cuaternión real, entonces $q + wQ$ es un bicuaternión. Los bicuaterniones de Clifford satisfacen la ley del producto de la multiplicación, pero la multiplicación no es asociativa. En su trabajo posterior, Clifford introdujo las álgebras que llevan su nombre. Hay álgebras de Clifford con unidades 1, $e_1,...,e_{n-1}$ tales que el cuadrado de cada $e_i = -1$ y $e_ie_j = -e_je_i$ para $i \neq j$. Cada producto de dos o más unidades es una nueva unidad y por lo mismo hay 2^n unidades diferentes. Todos los productos son asociativos. Una forma es un escalar multiplicado por una unidad y un álgebra es generada por la suma y el producto de las formas.

El flujo de nuevos sistemas de hipernúmeros continuó surgiendo y la variedad se hizo enorme. En un ensayo, «Linear Associative Algebra» («Algebra Lineal Asociativa») leído en 1870 y publicado en forma litográfica en 1871 [23], Benjamin Peirce (1809-1880), profesor de matemáticas en la universidad de Harvard, definió y proporcionó un resumen de las álgebras lineales asociativas ya conocidas en sus días. La palabra lineal significa que el producto de dos unidades primarias cualesquiera es reducido a una de las unidades como cuando i multiplicado por j se reemplaza por k en los cuaterniones y la palabra asociativa significa que la multiplicación es asociativa. La adición en estas álgebras tiene las propiedades comunes de los números reales y complejos. En este ensayo, Peirce introdujo la idea de un elemento nilpotente, esto es, un elemento A tal que $A^n = 0$ para algún entero positivo n, y un elemento idempotente, esto es, $A^n = 1$ para algún n. También demostró que un álgebra donde al menos un elemento no es nilpotente posee un elemento idempotente.

---

[21] 1853, p. 650.

[22] *Proc. Lond. Math. Soc.*, 4, 1873, 381-395 = *Coll. Math. Papers*, 181-200 y *Amer. Jour. of Math.*, 1, 1878, 350-358 = *Coll. Math. Papers*, 266-276.

[23] *Amer. Jour. of Math.*, *4*, 1881, 97-229.

La cuestión de por qué se concedía tanta libertad con tal variedad de tales álgebras también se les había ocurrido a los matemáticos durante el mismo período en que estaban creando álgebras específicas. Gauss era un convencido (*Werke*, 2, 178) de que una extensión de los números complejos que conservara las propiedades básicas de los números complejos resultaba imposible. Es significativo que cuando Hamilton buscó un álgebra tridimensional para representar los vectores en el espacio, conformándose con los cuaterniones carentes de conmutatividad, no pudo probar que no había una álgebra conmutativa tridimensional. Tampoco Grassmann tenía tal prueba.

Más tarde se formularon los teoremas precisos. En 1878 F. Georg Frobenius (1849-1917) [24] demostró que las únicas álgebras lineales asociativas con coeficientes reales (de las unidades primarias), con un número finito de unidades primarias, un elemento unidad para la multiplicación, y obedeciendo las leyes del producto son las de los números reales, los números complejos y los cuaterniones reales. El teorema lo demostró independientemente Charles Sanders Peirce (1839-1914) en un apéndice al artículo de su padre [25]. Otro resultado importante, al que Weierstrass había llegado en 1861, establece que las únicas álgebras lineales asociativas con coeficientes reales o complejos (de las unidades primarias), con un número finito de unidades primarias, obedeciendo la ley del producto y con multiplicación conmutativa, son las de los números reales y complejos. Dedekind obtuvo el mismo resultado alrededor de 1870. El resultado de Weierstrass fue publicado en 1884 [26] y el de Dedekind el año siguiente [27].

En 1898, Adolf Hurwitz (1859-1919) [28] demostró que los números reales, los números complejos, los cuaterniones reales y los bicuaterniones de Clifford son las únicas álgebras lineales asociativas que cumplen la ley del producto.

Estos teoremas son valiosos porque nos dicen lo que podemos esperar en las extensiones del sistema de los números complejos si deseamos conservar al menos algunas de sus propiedades algebraicas.

---

[24] *Jour. für Math.*, 84, 1878, 1-63 = *Ges. Abh.*, 1, 343-405.

[25] *Amer. Jour. of Math.*, 4, 1881, 225-229.

[26] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1884, 395-410 = *Math. Werke*, 2, 311-332.

[27] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1885, 141-159 y 1887, 1-7 = *Werke*, 2, 1-27.

[28] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1898, 309-316 = *Math. Werke*, 2, 565-571.

Si Hamilton hubiera conocido estos teoremas, se habría evitado años de labor en su búsqueda de un álgebra tridimensional vectorial.

El estudio de las álgebras lineales con un número finito, y aun infinito, de unidades generadoras (primarias) y con o sin división continuó siendo una materia activa bien entrado el siglo XX. Autores como Leonard Eugene Dickson y J. H. M. Wedderburn contribuyeron mucho a la teoría.

Bibliografía

Clifford, W. K.: *Collected Mathematical Papers* (1882), Chelsea (reimpresión), 1968.

Collins, Joseph V.: «An elementary exposition of Grassmann's Ausdehnungslehre», *Amer. Math. Month.* 6, 1899, varias partes, y 7, 1900, varias partes.

Coolidge, Julian L.: *A History of geometrical methods*, Dover (reimpresión), 1963, pp. 252-264.

Crowe, Michael J.: *A History of vector analysis*, University of Notre Dame Press, 1967.

Dickson, Leonard E.: *Linear Algebras*, Cambridge University Press, 1914.

Gibbs, Josiah, W. y Wilson, E. B.: *Vector analysis* (1901). Dover (reimpresión), 1960.

Grassmann, H. G.: *Die lineale Ausdehnungslehre* (1844), Chelsea (reimpresión), 1969.

— : *Gesammelte mathematische und physikalische Werke*, 3 vols., B. G. Teubner, 1894-1911; vol. 1, Parte I, 1-319 contiene *Die lineale Ausdehnungslehre;* vol. 1, Parte II, 1-383 contiene *Die Ausdehnungslehre.*

Graves, R. P.: *Life of Sir William Rowan Hamilton*, 3 vols., Longmans Green, 1882-1889.

Hamilton, sir William R.: *Elements of quaternions*, 2 vols., 1866, 2.ª ed., 1899-1901, Chelsea (reimpresión), 1969.

— : *Mathematical papers*, Cambridge University Press, 1967, vol. 3.

— : «Papers in memory of Sir William R. Hamilton», *Scripta Math.*, 1945; also in *Scripta Math.*, 10, 1944, 9-80.

Heaviside, Oliver: *Electromagnetic theory*, Dover (reimpresión), 1950, vol. 1.

Klein, Felix: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea (reimpresión), 1950, vol. 1, pp. 167-191; vol. 2 pp. 2-12.

Maxwell, James Clerk: *The Scientific Papers*, 2 vols., Dover (reimpresión), 1965.

Peacock, George: «Report on the recent progress and present state of certain branches of analysis», *British Assn. for Advancement of Science Report for 1833*, Londres, 1834.

— : *A treatise on algebra*, 2 vols., 2.ª ed., Cambridge University Press, 1845; Scripta Mathematica (reimpresión), 1940.

Shaw, James B.: *Synopsis of linear associative algebra*, Carnegie Institution of Washington, 1907.

Smith, David Eugene: *A source book in mathematics*, Dover (reimpresión), 1959, vol. 2, pp. 677-696.

Study, E.: «Theorie der gemeinen und höheren complexen Grössen», *Encyk. der Math. Wiss.*, B. G. Teubner, 1898, I, 147-183.

Capítulo 33
DETERMINANTES Y MATRICES

Tanta es la ventaja de un lenguaje bien construido que su notación simplificada a menudo se convierte en fuente de teorías profundas.

P. S. LAPLACE

1. Introducción

Los determinantes y matrices fueron objeto, durante el siglo XIX, de gran atención, que produjo miles de artículos y, sin embargo, no constituyeron una gran innovación en matemáticas. El concepto de vector, que desde el punto de vista matemático no es más que una colección de tripletas ordenadas, tiene un significado físico directo como una fuerza o una velocidad, y con esto se describe matemáticamente de inmediato lo que la física dice. Lo mismo se aplica con igual fuerza al gradiente, la divergencia y el rotacional. De esta forma, aunque matemáticamente dy/dx es sólo un símbolo para una expresión larga con el límite de $\Delta y/\Delta x$, la derivada es en sí misma un concepto poderoso que nos permite pensar directa y creativamente acerca de sucesos físicos. Así, aunque las matemáticas superficialmente consideradas no son más que un lenguaje o una abreviación, sus conceptos más vitales son aquellos que proporcionan claves para nuevas esferas del pensamiento. En contraste, los determinantes y las matrices representan únicamente innovaciones del lenguaje. Son expresiones estenográficas para ideas que ya existen en una forma más complicada. Por sí mismas no dicen directamente algo que no

esté ya dicho por las ecuaciones o transformaciones, aunque en formas mucho más amplias. Ni los determinantes ni las matrices han influido profundamente el curso de las matemáticas a pesar de su utilidad como expresiones compactas y de lo sugestivo de las matrices como expresiones compactas para la discusión de los teoremas generales de la teoría de grupos. Sin embargo, ambos conceptos han demostrado ser herramientas altamente útiles y son hoy día parte del aparato de las matemáticas.

2. Algunos nuevos usos de los determinantes

Los determinantes surgieron en la solución de sistemas de ecuaciones lineales (cap. 25, sec. 3). Este problema y la teoría de la eliminación, la transformación de coordenadas, el cambio de variables en las integrales múltiples, la solución de los sistemas de ecuaciones diferenciales que surgen en el movimiento planetario, la reducción de las formas cuadráticas en tres o más variables y los haces de formas (un haz es el conjunto $A + \lambda B$, donde A y B son formas específicas y λ es un parámetro) a formas canónicas, todos dieron lugar a varios tipos nuevos de determinantes. Este trabajo del siglo XIX se siguió directamente del trabajo de Cramer, Bezout, Vandermonde, Lagrange y Laplace.

La palabra determinante, usada por Gauss para el discriminante de la forma cuadrática $ax^2 + 2bxy + cy^2$, la aplicó Cauchy a los determinantes que ya habían aparecido en los trabajos del siglo XVIII. La disposición de los elementos en cuadrado y la notación de subíndices dobles también se le deben a él [1]. Así, un determinante de tercer orden se escribe como (las líneas verticales fueron introducidas por Cayley en 1841)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} . \qquad (1)$$

En este artículo Cauchy proporcionó el primer tratamiento sistemático y casi moderno de los determinantes. Uno de los principales

[1] *Jour. de l'Ecole Poly.*, 10, 1815, 29-112 = *Œuvres* (2), 1, 91-169.

resultados es el teorema de la multiplicación para determinantes. Lagrange [2] ya había dado este teorema para determinantes de tercer orden, pero dado que las filas de su determinante eran las coordenadas de los vértices de un tetraedro, no estaba motivado para generalizarlo. Con Cauchy el teorema general, expresado en notación moderna, establece que

$$|a_{ij}| \cdot |b_{ij}| = |c_{ij}|, \quad (2)$$

donde $|a_{ij}|$ y $|b_{ij}|$ representan los determinantes de orden n-ésimo y $c_{ij} = \Sigma\, a_{ik}b_{kj}$. Esto es, el término de la i-ésima fila y la j-ésima columna del producto es la suma de los productos de los elementos correspondientes en la i-ésima fila de $|a_{ij}|$ y la j-ésima columna de $|b_{ij}|$. Este teorema había sido enunciado por Jacques P. M. Binet (1786-1856) en 1812 [3], pero no demostrado satisfactoriamente. Cauchy proporcionó también un argumento mejorado y una prueba del teorema de Laplace de desarrollo de los determinantes (cap. 25, sec. 3).

Heinrich F. Scherk (1798-1885), en su *Mathematische Abhandlungen* (*Disertación Matemática,* 1825), aportó varias nuevas propiedades de los determinantes. Formuló las reglas para la adición de dos determinantes que tienen una columna o fila en común y para la multiplicación de un determinante por una constante. Asimismo, estableció que el determinante de un cuadro que tiene como fila una combinación de dos o más filas es cero, y que el valor de un determinante triangular (todos los elementos superiores o inferiores de la diagonal principal son cero) es el producto de los elementos sobre la diagonal principal.

James Joseph Sylvester (1814-1897) se distinguió como uno de los estudiosos empeñados, por un período de cincuenta años, en la teoría de determinantes. Después de ganar el segundo Wrangler en los exámenes matemáticos (tripos), por ser judío se le impidió enseñar matemáticas en la universidad de Cambridge. De 1841 a 1845 fue profesor de la universidad de Virginia. Más adelante regresó a Londres y trabajó como actuario y abogado en 1845 a 1855. Se le ofreció la posición relativamente baja de profesor en una academia militar en Woolwich (Inglaterra) y en ella estuvo hasta 1871. Algu-

[2] *Nouv. Mém. de l'Acad. de Berlin,* 1773, 85-128 = *Œuvres,* 3, 577-616.
[3] *Jour. de l'Ecole Poly.*, 9, 1813, 280-302.

nos años de actividades misceláneas fueron seguidos por su nombramiento de profesor en la universidad Johns Hopkins, donde, empezando en 1876, impartió cátedra en teoría de invariantes. Inició la investigación en matemáticas puras en los Estados Unidos y fundó el *American Journal of Mathematics.* En 1884 regresó a Inglaterra, y a la edad de setenta años se convirtió en profesor de la universidad de Oxford, posición que mantuvo hasta su muerte.

Sylvester fue una persona vivaz, sensible, estimulante, apasionada, y aun excitable. Sus discursos eran brillantes y ocurrentes y presentaba sus ideas con gran fogosidad y entusiasmo. En sus artículos utilizaba un lenguaje brillante. Introdujo mucha terminología nueva y bromeaba comparándose a sí mismo con Adán, quien había dado los nombres a las bestias y las plantas. Aunque él relacionó muchos campos diversos, tales como la mecánica y la teoría de los invariantes, no se prestaba a trabajar en teorías de una manera sistemática y detallada. De hecho, frecuentemente publicaba conjeturas, y aunque muchas de éstas eran brillantes otras eran incorrectas. Lamentándose, reconocía que sus amigos del continente «cumplimentaban sus poderes de adivinación a costa de su juicio». Sus principales contribuciones versaron sobre ideas combinatorias y abstracciones a partir de desarrollos más concretos.

Uno de los logros más importantes de Sylvester fue un método mejorado para eliminar las x de los polinomios de grado n-ésimo y m-ésimo, al cual llamó método dialítico [4]. Así, para eliminar x de las ecuaciones

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0$$
$$b_0x^2 + b_1x + b_2 = 0$$

formó el determinante de quinto orden

$$\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix}. \qquad (3)$$

[4] *Phil. Mag.*, 16, 1840, 132-135 y 21, 1842, 534-539 = *Coll. Math. Papers,* 1, 54-57 y 86-90.

La anulación de este determinante es la condición necesaria y suficiente para que las dos ecuaciones tengan una raíz común. Sylvester no proporcionó prueba alguna. El método lleva, como Cauchy mostró [5], a los mismos resultados que los métodos de Euler y Bezout.

La fórmula para la derivada de un determinante cuando los elementos son funciones de t fue dada por Jacobi en 1841 [6]. Si las a_{ij} son funciones de t, A_{ij} es el menor complementario de a_{ij}, y D es el determinante entonces

$$\frac{\partial D}{\partial a_{ij}} = A_{ij}$$

y

$$\frac{dD}{dt} = \sum_{i,j} A_{ij} a'_{ij},$$

donde las primas denotan la diferenciación con respecto a t.

Los determinantes fueron empleados en otro contexto, a saber, en el cambio de variables de una integral múltiple. Jacobi (1832 y 1833) encontró primero los resultados particulares. Más adelante, Eugene Charles Catalan (1814-1894), en 1839 [7], proporcionó el resultado que hoy en día es familiar a los estudiantes de cálculo. De ese modo, la integral doble

$$\iint F(x, y)\, dx\, dy \tag{4}$$

bajo el cambio de variables dado por

$$x = f(u, v), \qquad y = g(u, v) \tag{5}$$

se convierte en

$$\iint G(u, v) \begin{vmatrix} f_u & f_v \\ g_u & g_v \end{vmatrix} du\, dv, \tag{6}$$

[5] *Exercices d'analyse et de physique mathématique*, 1, 1840, 385-422 = *Œuvres* (2), 11, 466-509.

[6] *Jour. für Math.*, 22, 1841, 285-318 = *Werke*, 3, 355-392.

[7] *Mémoires couronnés par l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, 14, 1841.

donde $G(u.v) = F(x(u,v),y(u,v))$. El determinante en (6) es llamado el jacobiano de x e y con respecto a u y v, o el determinante funcional.

Jacobi dedicó un ensayo importante [8] a los determinantes funcionales. En este artículo considera n funciones $u_1,\ldots,u_n$ cada una de las cuales es una función de x_1, $x_2,\ldots,x_n$ y formula la cuestión de cuándo, de tales n funciones, pueden ser eliminadas las x_i de tal manera que las u_i estén relacionadas por una ecuación. Si esto no es posible, se dice que las funciones u_i son independientes. La respuesta es que si el jacobiano de las u_i con respecto a las x_i se anula, las funciones no son independientes e inversamente. También aporta el teorema del producto para jacobianos. Esto es, si las u_i son funciones de las y_i y las y_i son funciones de las x_i, entonces el jacobiano de las u_i con respecto a las x_i es el producto del jacobiano de las u_i con respecto a las y_i y el jacobiano de las y_i con respecto a las x_i.

3. Los determinantes y las formas cuadráticas

En el siglo XVIII ya se había introducido el método para resolver el problema de transformar las ecuaciones de las secciones cónicas y superficies cuádricas a formas más simples mediante la selección de ejes coordenados que tienen las direcciones de los ejes principales. La clasificación de las superficies cuádricas en términos de los signos de los términos de segundo grado cuando la ecuación se encuentra en forma canónica, esto es, cuando los ejes principales son los ejes coordenados, la aportó Cauchy en su *Leçons sur les applications du calcul infinitésimal à la géométrie* (*Lecciones sobre las aplicaciones del cálculo infinitesimal a la geometría*, 1826) [9]. Sin embargo, no estaba claro que resulte siempre el mismo número de términos positivos y negativos en la reducción a la forma canónica. Sylvester contestó esta pregunta con su ley de inercia de formas cuadráticas en n variables [10]. Ya se sabía que

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j \tag{7}$$

[8] *Jour. für Math.*, 22, 1841, 319-359 = *Werke*, 3, 393-438.
[9] *Œuvres* (2), 5, 244-285.
[10] *Phil. Mag.* (4), 4, 1852, 138-42 = *Coll. Math. Papers*, 1, 378-381.

siempre puede ser reducido a la suma de r cuadrados [11]

$$y_1^2 + \dots + y_s^2 - y_{s+1}^2 - \dots - y_{r-2}^2 \tag{8}$$

mediante una transformación real lineal

$$x_i = \sum_j b_{ij}y_j, \qquad i = 1, 2, \dots, n,$$

con un determinante que no se anula. La ley de Sylvester establece que el número s de términos positivos y $r - s$ de negativos es siempre el mismo, cualquiera que sea la transformación real usada. Considerando la ley como autoevidente, Sylvester no proporcionó prueba alguna.

La ley fue redescubierta y probada por Jacobi [12]. Si una forma es positiva o cero para todos los valores reales de las variables, se la llama definida positiva. Entonces todos los signos de (8) son positivos, y $r = n$. Es llamada semidefinida si puede tomar valores positivos o negativos (en cuyo caso $r < n$), y definida negativa cuando la forma es siempre negativa o cero (y $r = n$). Estos términos los introdujo Gauss en su *Disquisitiones Arithmeticae* (sec. 271).

El estudio posterior de la reducción de formas cuadráticas supone la noción de ecuación característica de una forma cuadrática o de un determinante. Una forma cuadrática en tres variables se escribía, comúnmente, durante el siglo XVIII y la primera parte del XIX, como

$$Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz$$

y en tiempos más recientes como

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{22}x_2x_3. \tag{9}$$

En la última notación la forma era asociada con el determinante

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \qquad a_{ij} = a_{ji}. \tag{10}$$

---

[11] r es el rango de la forma, esto es, el rango de la matriz de los coeficientes. Para la noción de rango véase la sec. 4.

[12] *Jour. für. Math.*, 53, 1857, 265-270 = *Werke*, 3, 583-590; véase también 593-598.

La ecuación característica o ecuación latente de la forma o el determinante es

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0, \tag{11}$$

y los valores de λ que satisfacen esta ecuación son llamados raíces características o raíces latentes. A partir de estos valores de λ son obtenidas fácilmente las longitudes de los ejes principales [13].

La noción de ecuación característica aparece implícitamente en el trabajo de Euler [14] sobre la reducción de las formas cuadráticas en tres variables a sus ejes principales, aunque haya fallado en probar la «realidad» de las raíces características. La noción de ecuación característica aparece explícitamente en el trabajo de Lagrange sobre sistemas de ecuaciones diferenciales lineales [15], y en el trabajo de Laplace en la misma área [16].

Lagrange, al tratar con el sistema de ecuaciones diferenciales para el movimiento de seis planetas en su época, estaba interesado en las perturbaciones seculares (de largos períodos) que estos éstos ejercían sobre los otros. Su ecuación característica (también llamada ecuación secular) era para un determinante de orden seis y los valores de λ determinaban soluciones del sistema. Fue capaz de descomponer la ecuación de sexto grado y obtener información acerca de las raíces. Laplace demostró en su *Mécanique céleste* que si todos los planetas se mueven en la misma dirección, entonces las seis raíces características son reales y distintas. La «realidad» de los valores característicos para el problema cuadrático en tres variables fue establecido por Hachette, Monge y Poisson [17].

---

[13] Por una transformación lineal $x_i' = \Sigma m_{ij} x_j$, $i,j = 1,2,3$, la forma (9) puede ser reducida a $\Sigma_{j=1}^{3} \lambda_j x_j'^2$, donde las λ_j son las raíces características de (11). En el lenguaje de las matrices, la matriz M de la transformación es ortogonal; esto es, la traspuesta de M es igual a la inversa de M.

[14] El cap. 5 del Apéndice en su *Introductio* (1748) = *Opera* (1), 9, 379-392.

[15] *Misc. Taur.*, 3, 1762-1765 = *Œuvres*, 1, 520-534 y *Mém. de l'Acad. des Sci., Paris*, 1774 = *Œuvres*, 6, 655-666.

[16] *Mém. de l'Acad. des Sci., Paris*, 1772, pub. 1775 = *Œuvres*, 8, 325-366.

[17] Hachette y Monge: *Jour. de l'Ecole Poly.*, 4, 1801-1802, 143-169; Poisson y Hachette, *ibid.*, 170-172.

Cauchy reconoció el problema del valor característico común en la obra de Euler, Lagrange y Laplace. En sus *Leçons* de 1826 [18] tomó el problema de la reducción de la forma cuadrática en tres variables y demostró que la ecuación característica es invariante para cualquier cambio de los ejes rectangulares. Tres años más tarde, en sus *Exercices de mathématiques (Ejercicios de matemáticas)* [19], atacó el problema de las desigualdades seculares de las trayectorias planetarias. En el transcurso de su trabajo demostró que dos formas cuadráticas en n variables

$$A = \sum_{i,j=1}^{n} a_{ij}x_ix_j, \qquad B = \sum_{i,j=1}^{n} b_{ij}x_ix_j$$

(la B de Cauchy era una suma de cuadrados) podían ser reducidas simultáneamente por una transformación lineal

$$x_1 = c_{11}x_1' + \ldots + c_{1n}x_n'$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_n = c_{n1}x_1' + \ldots + c_{nn}x_n'$$

a una suma de cuadrados. También resolvió el problema de encontrar los ejes principales para formas en cualquier número de variables y en este trabajo usó también la noción de raíces características.

Su trabajo se reduce a lo siguiente: si A y B son dos formas cuadráticas cualesquiera dadas, entonces se puede considerar la familia *(Schaar)* de formas $uA + vB$, donde u y v son parámetros arbitrarios. Las raíces latentes de la familia son los valores de la proporción $-u/v$ para los que el determinante de la familia $|ua_{ij} + vb_{ij}|$ es cero. Cauchy demostró que las raíces latentes son todas reales en el caso especial de que una de las formas es definida positiva para todos los valores de las variables reales distintos de cero. Ya que el determinante de $uA + vB$ es simétrico ($d_{ij} = d_{ji}$) y B podía ser el determinante identidad ($b_{ij} = 0$ para toda $i \neq j$ y $b_{ii} = 1$), Cauchy demostró que cualquier determinante simétrico real de cualquier orden tiene raíces características reales. Los resultados de Cauchy, duplicados por los de Jacobi en 1834 [20], excluían la raí-

[18] *Œuvres* (2), 5, 244-285.

[19] 4, 1829, 140-160 = *Œuvres* (2), 9, 174-195.

[20] *Jour. für Math.*, 12, 1834, 1-69 = *Werke*, 3, 191-268.

ces latentes iguales. El término ecuación característica se debe a Cauchy [21].

La noción de determinantes semejantes también surgió del trabajo sobre transformaciones. Dos determinantes A y B son semejantes si existe un determinante P distinto de cero tal que $A = P^{-1}BP$. Las transformaciones semejantes fueron consideradas por Cauchy, quien mostró en sus *Leçons* de 1826 [22] que tienen los mismos valores característicos. La importancia de las transformaciones semejantes radica en clasificar transformaciones proyectivas (cap. 38, sec. 5), un problema que por largo tiempo fue tratado sintéticamente. Si una figura F es relacionada con una figura G por una transformación lineal A y si otra transformación B transforma F en F' y G en G', la transformación C que lleva F' en G' tendrá las mismas propiedades que A. La transformación es $C = BAB^{-1}$, ya que B^{-1} lleva F' en F, A lleva F en G, y B lleva G en G'.

En 1858, Weierstrass [23] proporcionó un método general para reducir dos formas cuadráticas simultáneamente a sumas de cuadrados. También probó que si una de las formas es definida positiva, la reducción es posible aun cuando algunas de las raíces latentes sean iguales. El interés de Weierstrass en este problema surgió de un problema dinámico de pequeñas oscilaciones cerca de una posición de equilibrio, y demostró, por medio de su trabajo sobre formas cuadráticas, que la estabilidad no es destruida por la presencia de períodos iguales en el sistema, contrariamente a las suposiciones de Lagrange y Laplace.

En 1851 [24], Sylvester, trabajando en el contacto y la intersección de curvas y superficies de segundo grado, fue llevado a considerar la clasificación de familias de tales cónicas y superficies cuadráticas. En particular, buscó la forma canónica de cualquier familia. Escribiendo una familia en la forma $A + \lambda B$ donde

$$A = ax^2 + by^2 + cz^2 + 2dyz + 2ezx + 2fxy$$
$$B = Ax^2 + By^2 + Cz^2 + 2Dyz + 2Ezx + 3Fxy$$

---

[21] *Exercices d'analyse et de physique mathématique*, 1, 1840, 53 = *Œuvres* (2), 11, 76.

[22] *Œuvres* (2), 5, 244-285.

[23] *Monatsber. Berliner Akad.*, 1858, 207-220 = *Werke*, 1, 233-246.

[24] *Phil. Mag.* (4), 1, 1851, 119-140 = *Coll. Math. Papers*, 1, 219-240.

consideró el determinante

$$\begin{vmatrix} a + \lambda A & f + \lambda F & e + \lambda E \\ f + \lambda F & b + \lambda B & d + \lambda D \\ e + \lambda E & d + \lambda D & c + \lambda C \end{vmatrix}. \qquad (12)$$

Su método de clasificación introdujo la noción de divisor elemental. Los elementos del determinante de $A + \lambda B$ son polinomios en λ. Sylvester demostró que si todos los menores de cualquier orden de $|A + \lambda B|$ tienen un factor $\lambda + \varepsilon$ en común, este factor continuará siendo común al mismo sistema de menores cuando A y B son transformados simultáneamente por una transformación lineal de sus variables. También demostró que si todos los menores de orden i-ésimo tienen un factor $(\lambda + \varepsilon)^{a}$ los menores de orden $(i + h)$ contienen el factor $(\lambda + \varepsilon)^{(h+1)\,a}$. Los diversos factores lineales a la potencia con que aparecen en el máximo cómun divisor $D_i(\lambda)$ de los menores de i-ésimo orden para cada i son los *divisores elementales* de $|A + \lambda B|$ o de cualquier determinante general. Los cocientes de $D_i(\lambda)$ por $D_{i-1}(\lambda)$ para cada i son llamados factores invariantes de $|A + \lambda B|$. Sylvester no demostró que los factores invariantes constituyen un conjunto completo de invariantes para las dos formas cuadráticas.

Weierstrass[25] completó la teoría de las formas cuadráticas y extendió la teoría de las formas bilineales; una forma bilineal es

$$a_{11}x_1y_1 + a_{12}x_1y_2 + \ldots + a_{nn}x_ny_n.$$

Usando la notación de los divisores elementales de Sylvester, Weierstrass obtuvo la forma canónica de la familia $A + \lambda B$, donde A y B no son necesariamente simétricas, pero sí sujetas a la condición de que $|A + \lambda B|$ no es idénticamente cero. Además, demostró el inverso de un teorema debido a Sylvester. El inverso establece que si el determinante de $A + \lambda B$ coincide en sus divisores elementales con el determinante de $A' + \lambda B'$, entonces se puede encontrar un par de transformaciones lineales que transforman simultáneamente A en A' y B en B'.

Entre la multitud de teoremas sobre determinantes están los que se dedican a la solución de m ecuaciones lineales en n incógnitas,

[25] *Monatsber. Berliner Akad.*, 1868, 310-338 = *Werke*, 2, 19-44.

Henry J. S. Smith (1826-1883) [26] introdujo los términos menor aumentado (o ampliado) y no ampliado para discutir la existencia y número de soluciones de, por ejemplo, un sistema como

$$\begin{aligned} a_1x + a_2y &= f \\ b_1x + b_2y &= g \\ c_1x + c_2y &= h. \end{aligned}$$

Los menores ampliados o no son

$$\begin{Vmatrix} a_1 & a_2 & f \\ b_1 & b_2 & g \\ c_1 & c_2 & h \end{Vmatrix} \quad \text{y} \quad \begin{Vmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{Vmatrix}.$$

Una serie de resultados debida a muchos autores, incluyendo a Kronecker y Cayley, condujo a los resultados generales ahora comúmente establecidos en términos de matrices, pero que en mitad del siglo XIX y fueron establecidos en términos de determinantes ampliados o no. Los resultados generales sobre m ecuaciones con n incógnitas con m mayor que, igual a, o menor que n —las ecuaciones pueden ser homogéneas (los términos constantes son cero) o no homogéneas— están enunciados, por ejemplo, en el libro de Charles L. Dodgson (Lewis Carroll, 1832-1898) *An elementary theory of Determinants* (*Teoría elemental de determinantes*, 1867). En textos posteriores aparece esta condición: para que un conjunto de n ecuaciones lineales no homogéneas en n incógnitas pueda ser compatible es necesario y suficiente que el máximo orden de los menores no nulos sea del mismo orden en los menores ampliados o no, o en términos del lenguaje matricial, que los rangos de las dos matrices sean el mismo.

Se obtuvieron nuevos resultados sobre determinantes a lo largo del siglo XIX. Como una ilustración tenemos un teorema demostrado por Hadamard en 1893 [27], aunque conocido y demostrado por muchos otros antes y después de esta fecha. Si los elementos del determinate $D = |a_{ij}|$ satisfacen la condición de que $|d_{ij}| \leqslant A$ entonces

$$|D| \leq A^n \cdot n^{n/2}.$$

[26] *Phil. Trans.*, 151, 293-326 = *Coll. Math. Papers*, 1, 367-409.

[27] *Bull. des Sci. Math.* (2), 17, 1893, 240-246 = *Œuvres*, 1, 239-245.

Los teoremas anteriores sobre determinantes sólo representan un pequeño ejemplo de la multitud de los que han sido establecidos. Además de una gran variedad de otros teoremas sobre determinantes generales, hay cientos de otros sobre determinantes de formas especiales tales como los determinantes simétricos ($a_{ij} = a_{ji}$), determinantes antisimétricos ($a_{ij} = - a_{ji}$), ortogonantes (determinantes de transformaciones coordenadas ortogonales), determinantes orlados (determinantes extendidos por la adición de filas y columnas), determinantes compuestos (los elementos son determinantes ellos mismos), y muchos otros tipos especiales.

4. Matrices

Diríamos que el campo de las matrices estuvo bien desarrollado aun antes de crearse. Los determinantes, como sabemos, fueron estudiados a partir de mediados del siglo XVIII. Un determinante contiene un cuadro de números y por lo general interesa el valor del cuadro, dado por la definición del determinante. Parecía deducirse de la inmensa cantidad de trabajo sobre los determinantes que el cuadro podía ser estudiado en sí mismo y manipulado para muchos propósitos, fuera o no el valor del determinante lo interesante. Quedaba por reconocerse que al cuadro como tal se le podía proporcionar una identidad independiente de la del determinante. El cuadro por sí mismo es llamado matriz. La palabra *matriz* fue usada por primera vez por Sylvester [28] cuando en realidad deseaba referirse a un cuadro rectangular de números y no podía usar la palabra determinante, aunque en aquella ocasión sólo le interesaban los determinantes que podía formar a partir de los elementos del cuadro rectangular. Más tarde, como ya hemos hecho notar en la sección precedente, los cuadros ampliados se emplearon con libertad sin hacer mención alguna de las matrices. Las propiedades básicas de las matrices, como veremos, fueron también establecidas en el desarrollo de los determinantes.

Es cierto, como dice Arthur Cayley (en el artículo de 1855 al que nos referiremos más abajo), que lógicamente la idea de matriz precede a la de determinante, pero históricamente el orden fue el inverso y esto se debe a que las propiedades básicas de las matrices

[28] *Phil. Mag.* (3), 37, 1850, 363-370 = *Coll. Math. Papers,* 1, 145-151.

ya estaban claras cuando fueron introducidas las matrices. Así, la impresión general entre los matemáticos referente a que las matrices representaban una creación altamente original e independiente inventada por los matemáticos puros cuando adivinaron la utilidad potencial de la idea es erróneo. Porque el empleo de las matrices estaba bien establecido se le ocurrió a Cayley introducirlas como entidades diferentes. Dice: «ciertamente no obtuve la noción de matriz de alguna manera a partir de los cuaterniones; fue directamente de la de determinante o como una forma conveniente de expresar las ecuaciones:

$$x' = ax + by\,''$$
$$y' = cx + dy.$$

Y así introdujo la matriz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

que representa la información esencial acerca de la transformación. Como Cayley fue el primero en aislar la matriz por sí misma y el primero en publicar una serie de artículos sobre ellas, se le aceptó generalmente como el creador de la teoría de matrices.

Cayley, nacido en 1821, de una antigua y talentosa familia inglesa, mostró habilidad matemática en la escuela. Sus profesores convencieron a su padre para enviarlo a Cambridge, en vez de ponerlo en los negocios de la familia. En Cambridge fue el decano de los «wrangler» en los exámenes de matemáticas (tripos) y ganó el premio Smith. Lo seleccionaron como un *Fellow* del Trinity College en Cambridge y tutor asistente, pero después de tres años la abandonó porque deseaba tomar las órdenes sagradas. Volvió a las leyes y dedicó los quince años siguientes a esa profesión. Durante este período fue capaz de emplear un tiempo considerable en las matemáticas y publicó cerca de 200 artículos. Y durante este período empezó también su larga amistad, y colaboración, con Sylvester.

En 1863 obtiene el nombramiento para la recién creada cátedra Sadleriana de matemáticas en Cambridge. Excepto por el año 1882, en que estuvo en la universidad Johns Hopkins como invitado de Sylvester, permaneció en Cambridge hasta su muerte, acaecida en 1895. Fue un escritor muy prolífico y creador de varias áreas: la geometría analítica de n dimensiones, la teoría de los determinantes,

transformaciones lineales, superficies alabeadas y teoría de matrices. Junto con Sylvester, fundó la teoría de invariantes. Por estas contribuciones tan numerosas recibió muchos honores.

Contrariamente a Sylvester, Cayley fue un hombre de temperamento equilibrado, juicio sobrio y serenidad, generoso en ayudar y apoyar a otros. Además de su trabajo excelente en el derecho y prodigiosos logros en matemáticas, encontró tiempo para su interés en la literatura, pintura, arquitectura y viajes.

Fue en relación con el estudio de los invariantes bajo trasnformaciones lineales (cap. 39, sec. 2) como Cayley introdujo por primera vez las matrices para simplificar la notación [29]. Aquí proporcionó algunas nociones básicas. Esto fue seguido por su principal contribución a la materia, «A memoir on the theory of matrices» («Una memoria sobre la teoría de las matrices») [30].

Por brevedad daremos las definiciones de Cayley para matrices 2 por 2 y 3 por 3, a pesar de que las definiciones son válidas para matrices n por n y en algunos casos para matrices rectangulares. Dos matrices son iguales si sus elementos correspondientes son iguales. Cayley define la matriz cero y la matriz unidad como

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La suma de dos matrices es definida como la matriz cuyos elementos son la suma de los elementos correspondientes de los dos sumandos. Señala que esta definición se aplica a dos matrices cualesquiera m por n y la adición es asociativa y convertible (conmutativa). Si m es un escalar y A una matriz, entonces mA está definida como la matriz cuyos elementos son cada uno m veces los correspondientes elementos de A.

Cayley toma directamente de la representación del efecto de dos transformaciones sucesivas la definición de multiplicación de dos matrices. Así, si la transformación

$$x' = a_{11}x + a_{12}y$$
$$y' = a_{21}x + a_{22}y$$

---

[29] *Jour. für Math.*, 50, 1855, 282-285 = *Coll. Math. Papers*, 2, 185-188.
[30] *Phil. Trans.*, 148, 1858, 17-37 = *Coll. Math. Papers*, 2, 457-496.

es seguida por la transformación

$$x'' = b_{11}x' + b_{12}y'$$
$$y'' = b_{21}x' + b_{12}y'$$

entonces la relación entre x'', y'' y x, y está dada por

$$x'' = (b_{11}a_{11} + b_{12}a_{21})x + (b_{11}a_{12} + b_{12}a_{22})y$$
$$y'' = (b_{21}a_{11} + b_{22}a_{21})x + (b_{21}a_{12} + b_{22}a_{22})y.$$

De aquí Cayley definió el producto de las matrices como

$$\begin{pmatrix} b_{11}b_{12} \\ b_{21}b_{22} \end{pmatrix} \begin{pmatrix} a_{11}a_{12} \\ a_{21}a_{22} \end{pmatrix} = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix}.$$

Esto es, el elemento c_{ij} en el producto es la suma de los productos de los elementos en la i-ésima fila del factor izquierdo y los elementos correspondientes de la j-ésima columna del factor derecho. La multiplicación es asociativa, pero no es conmutativa. Cayley señala que una matriz m por n puede ser compuesta únicamente con una matriz n por p.

En este mismo artículo, establece que el inverso de

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix} \quad \text{está dada por} \quad \frac{1}{\nabla}\begin{pmatrix} \partial_a\nabla, & \partial_{a'}\nabla, & \partial_{a''}\nabla \\ \partial_b\nabla, & \partial_{b'}\nabla, & \partial_{b''}\nabla \\ \partial_c\nabla, & \partial_{c'}\nabla, & \partial_{c''}\nabla \end{pmatrix},$$

donde ∇ es el determinante de la matriz y $\partial_x\Delta$ es el cofactor de x en este determinante, esto es, el menor de x con el signo oportuno. El producto de una matriz y su inverso es la matriz unidad, denotada por I.

Cuando $\nabla = 0$, la matriz es indeterminada (en terminología moderna, singular) y no tiene inversa. Cayley aseguró que el producto de dos matrices puede ser cero sin ser alguna de ellas necesariamente cero, si cualquiera de ellas es indeterminada. En realidad, Cayley estaba equivocado, han de ser indeterminadas las dos. Porque si $AB = 0$, $A \neq 0$, $B \neq 0$ y sólo A es indeterminada, entonces la inversa B^{-1} de B existe, y $ABB^{-1} = 0$. $B^{-1} = 0$. Pero $BB^{-1} = I$. Por tanto, $AI = 0$ y $A = 0$.

La matriz transversa (traspuesta o conjugada) se define como aquella obtenida cambiando filas y columnas. Se argumenta (sin prueba) que $(LMN)' = N'M'L'$ donde la prima denota traspuesta. Si $M' = M$ entonces M se llama simétrica, y si $M' = -M$ entonces M es antisimétrica (o alternada). Cualquier matriz puede ser expresada como la suma de una matriz simétrica y una antisimétrica.

Otro concepto tomado de la teoría de determinantes es la ecuación característica de una matriz cuadrada. Para la matriz M está definida como

$$|M - xI| = 0,$$

donde $|M - xI|$ es el determinante de la matriz $M - xI$ e I es la matriz unitaria. Así si

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la ecuación característica (Cayley no usa el término, aunque Cauchy lo introdujo para los determinantes [sec. 3]) es

$$x^2 - (a + d)x + ad - bc = 0. \tag{13}$$

Las raíces de esta ecuación son las raíces características (valores propios) de la matriz.

En el artículo de 1858, Cayley anunció lo que ahora es conocido como el teorema de Cayley-Hamilton para matrices cuadradas de cualquier orden. El teorema dice que si x es sustituida por M en (13), la matriz resultante es la matriz cero. Cayley afirma que él ha verificado este teorema para el caso de 3 por 3 y que no se requiere más prueba. La asociación de Hamilton con este teorema se apoya en el hecho de que al introducir en sus *Lecciones sobre cuaternios* [31] la noción de una función vectorial lineal r' de otro vector r, aparece una transformación lineal de x, y y z en x', y' y z'. Demostró que la matriz de esta transformación satisface la ecuación característica de esa matriz, aunque no pensó formalmente en términos de matrices.

Otros matemáticos encontraron propiedades particulares de las raíces características de algunas clases de matrices. Hermite [32], de-

---

[31] 1853, p. 566.

[32] *Comp. Rend.*, 41, 1855, 181-183 = *Œuvres*, 1, 479-481.

mostró que si la matriz $M = M^*$, donde M^* es la traspuesta de la matriz formada al reemplazar cada elemento de M por su conjugado complejo (tales M son llamadas ahora hermitianas), entonces las raíces características son reales. En 1861, Clebsch [33] dedujo del teorema de Hermite que las raíces características diferentes de cero de una matriz antisimétrica real son imaginarias puras. Luego Arthur Buchheim [34] (1859-1888), demostró que si M es simétrica y si los elementos son reales, las raíces características son reales, aunque este resultado ya había sido establecido para los determinantes por Cauchy [35]. Henry Taber (1860-?), en otro ensayo [36], afirmó como evidente que si

$$x^n - m_1x^{n-1} + m_2x^{n-2} - \dots \pm m_n = 0$$

es la ecuación característica de cualquier matriz cuadrada M, entonces el determinante M es m_n, y si entendemos por menor principal de una matriz el *determinante* de un menor cuya diagonal es parte de la diagonal principal de M, entonces m_i es la suma de los i menores principales. En particular, entonces, m_1, que es también la suma de las raíces características, es la suma de los elementos de la diagonal principal. Esta suma es llamada traza de la matriz. Las demostraciones de las afirmaciones de Taber fueron proporcionadas por William Henry Metzler (1863-?) [37].

Frobenius [38] formuló una pregunta relacionada con la ecuación característica. Buscó el polinomio mínimo —el polinomio de menor grado— que satisface la matriz. Estableció que es el formado a partir de los factores del polinomio característico y es único. No fue hasta 1904 [39] cuando Kurt Hensel (1861-1941) demostró la unicidad de la afirmación de Frobenius. En el mismo artículo, Hensel también demostró que si $f(x)$ es un polinomio mínimo de una matriz M y $g(x)$ es cualquier otro polinomio satisfecho por M, entonces $f(x)$ divide a $g(x)$.

Frobenius [40] introdujo la noción de rango de una matriz en 1879,

---

[33] *Jour. für Math.*, 62, 1863, 232-245.
[34] *Messenger of Math.* (2), 14, 1885, 143-144.
[35] *Œuvres* (2), 9, 174-191.
[36] *Amer. Jour. of Math.*, 12, 1890, 337-396.
[37] *Amer. Jour. of Math.*, 14, 1891/1892, 326-377.
[38] *Jour. für Math.*, 84, 1878, 1-63 = *Ges. Abh.*, 1, 343-405.
[39] *Jour. für Math.*, 127, 1904, 116-166.
[40] *Jour. für Math.*, 86, 1879, 146-208 = *Ges. Abh.*, 1, 482-544.

aunque en relación con los determinantes. Una matriz A con m filas y n columnas (de orden m por n) tiene menores con k filas de todos los órdenes desde 1 (los elementos de A mismos) hasta el menor de los dos enteros m y n inclusive. Una matriz tiene rango r si y solamente si tiene al menos un menor de orden r cuyo determinante no es cero mientras que los determinantes de todos los menores de orden superior a r son cero.

Dos matrices A y B pueden ser relacionadas de diversas maneras. Son equivalentes si existen dos matrices no singulares U y V tales que $A = UBV$. Sylvester había demostrado en su trabajo sobre determinantes [41] que el máximo común divisor d_i^1 de los menores con i filas de B es igual al máximo común divisor d_i de los menores con i filas de A. Entonces H. J. S. Smith, trabajando con matrices con elementos enteros, demostró [42] que cada matriz A de rango p es equivalente a una matriz diagonal con elementos $h_1, h_2, \ldots, h_p$ debajo de la diagonal principal y tal que h_i divide a $h_i + 1$. Los cocientes $h_1 = d_1$, $h_2 = d_2/d_1, \ldots$ son llamados los factores invariantes de A. Además, si

$$h_i = p_1^{l_{i1}} p_2^{l_{i2}} \ldots p_k^{l_{ik}}$$

(donde los p_i son primos) estas varias potencias $p_j^{l_{ij}}$ son los divisores elementales de A. Los factores invariantes determinan los divisores elementales e inversamente.

Estas ideas sobre factores invariantes y divisores elementales, que surgen del trabajo de Sylvester y Weierstrass sobre determinantes (como se señaló anteriormente), fueron trasladadas a las matrices por Frobenius en su ensayo de 1878. El significado de estos factores invariantes y divisores elementales para matrices es que la matriz A es equivalente a la matriz B si y solamente si A y B tienen los mismos divisores elementales o factores invariantes.

Frobenius realizó más trabajo con los factores invariantes en su artículo de 1878 y organizó la teoría de los factores invariantes y divisores elementales en forma lógica [43]. El trabajo en el artículo de 1878 le permitió a Frobenius dar la primera demostración general del teorema de Cayley-Hamilton y modificar el teorema cuando

---

[41] *Phil. Mag.* (4), 1, 1851, 119-140 = *Coll. Math. Papers*, 1, 219-240.

[42] *Phil. Trans.*, 141, 1861-1862, 293-326 = *Coll. Math. Papers*, 1, 367-409.

[43] *Sitzungsber. Akad. Wiss. zu Berlin*, 1894, 31-44 = *Ges. Abh.*, 1, 577-590.

algunas de las raíces latentes (raíces características) de la matriz son iguales. En este ensayo demostró que cuando $AB^{-1} = B^{-1}A$, en cuyo caso existe un cociente bien definido A/B, entonces $(A/B)^{-1} = B/A$ y que $(A^{-1})^T = (A^T)^{-1}$, donde A^T es la traspuesta de A.

Las matrices ortogonales han recibido considerable atención. Aunque el término haya sido usado por Hermite en 1854 [44], la definición formal no es publicada hasta 1878 por Frobenius (véase referencia anterior). Una matriz M es ortogonal si es igual a la inversa de la traspuesta, esto es, si $M = (M^T)^{-1}$. Además de la definición, Frobenius demostró que si S es una matriz simétrica y T una matriz antisimétrica, una matriz ortogonal puede ser escrita en la forma $(S - T)/(S + T)$ o más simplemente $(I - T)/(I + T)$.

La noción de matrices semejantes, como muchos otros conceptos de la teoría de matrices, vino del trabajo anterior sobre determinantes, tan alejado en el tiempo como el de Cauchy. Dos matrices cuadradas A y B son semejantes si existe una matriz no singular P tal que $B = P^{-1}AP$. Las ecuaciones características de dos matrices semejantes A y B son la misma y las matrices tienen los mismos factores invariantes y los mismos divisores elementales. Para matrices con elementos complejos, Weierstrass demostró este resultado en su ensayo de 1868 (aunque él trabajó con determinantes). Ya que una matriz representa una transformación lineal homogénea, matrices similares pueden ser pensadas como si representaran la misma transformación pero referidas a dos sistemas coordenados diferentes.

Usando el concepto de matrices similares y de ecuación característica, Jordan [45] demostró que una matriz puede ser transformada a una forma canónica. Si la ecuación característica de una matriz J es

$$f(\lambda) = \lambda^n + b_1\lambda^{n-1} + \dots + b_n = 0$$

y si

$$f(\lambda) = (\lambda - \lambda_1)^{l_1} \dots (\lambda - \lambda_k)^{l_k},$$

donde las λ_i son diferentes, entonces sea

[44] *Cambridge and Dublin Math. Jour.*, 9, 1854, 63-67 = *Œuvres*, 1, 290-295.
[45] *Traité des substitutions*, 1870, libro II, 88-249.

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

que denota una matriz de orden l_i-ésimo. Jordan demostró que J puede ser transformada en una matriz similar teniendo la forma

$$\begin{pmatrix} J_1 & 0 & 0 \dots & 0 \\ 0 & J_2 & 0 \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots & J_k \end{pmatrix}.$$

Esta es la forma canónica, o normal, de Jordan de una matriz.

Frobenius también trató, en su ensayo de 1878, la transformación de semejanza de A en B, bajo el nombre de transformaciones contragredientes. En la misma discusión trató la noción de matrices congruentes o transformaciones cogredientes. Esto nos dice que si $A = P^T BP$ entonces A es congruente con B, escrito $A \stackrel{c}{=} B$. Por ejemplo, la transformación de la matriz A que resulta del intercambio simultáneo de las mismas filas y columnas de A es una transformación congruente. De igual manera, una matriz simétrica A de rango r puede ser reducida por medio de una transformación congruente a una matriz diagonal del mismo rango, esto es:

$$P^T AP = \begin{bmatrix} d_{11} & 0 & \dots & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{rr} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

Hay muchos teoremas básicos sobre matrices congruentes. Por ejemplo, si S es simétrica y S_1 es congruente con S, entonces S_1 es simétrica y si S es antisimétrica entonces S_1 es antisimétrica.

En su artículo de 1892 del *American Journal of Mathematics*,

Metzler introdujo las funciones trascendentes de una matriz, escribiendo cada una como una serie de potencias de una matriz. Dio las series para e^M, e^{-M}, log M, sen M y sen^{-1} M. Así

$$e^M = \sum_{n=0}^{\infty} M^n/n!.$$

Las ramificaciones de la teoría de matrices son numerosas. Las matrices han sido usadas para representar formas cuadráticas y bilineales. La reducción de tales formas a formas canónicas simples constituye el centro del trabajo sobre los invariantes de las matrices, conectados íntimamente con los hipernúmeros; Cayley, en su ensayo de 1858, desarrolló la idea de tratar los hipernúmeros como matrices.

Tanto los determinantes como las matrices han sido extendidos al caso de orden infinito. Los determinantes infinitos estaban involucrados en el trabajo de Fourier para la determinación de los coeficientes de una serie de Fourier de una función (cap. 28, sec. 2) y en el trabajo de Hill sobre la solución de ecuaciones diferenciales ordinarias (cap. 29, sec. 7). Artículos aislados sobre determinantes infinitos se escribieron entre estos dos grandes investigadores del siglo XIX, pero la mayor actividad es posterior a Hill.

Las matrices infinitas estaban incluidas implícita y explícitamente en el trabajo de Fourier, Hill y Poincaré, quien completó el trabajo de Hill. Sin embargo, el gran impulso aplicado al estudio de las matrices infinitas vino de la teoría de las ecuaciones integrales. (cap. 45). No podemos dedicar aquí espacio a la teoría de los determinantes y matrices de orden infinito [46].

En el trabajo básico sobre matrices, los elementos son números reales ordinarios aunque una gran parte de lo que se ha hecho en nombre de la teoría de números estaba circunscrito a elementos enteros. Sin embargo, éstos pueden ser números complejos y, por supuesto, casi cualquier otra cantidad. Naturalmente, las propiedades que las matrices poseen dependen de las propiedades de los elementos. Espacio significativo de la investigación de la última parte del siglo XIX y el inicio del XX se dedica a las propiedades de las matrices cuyos elementos están en un cuerpo abstracto. La importancia

[46] Véase la referencia a Bernkopf en la bibliografía al final del capítulo.

de la teoría de matrices en el instrumental del XIX de la física moderna no puede ser tratada aquí, pero en relación con ello es de interés un argumento profético hecho por Tait: «Cayley está forjando las armas para las futuras generaciones de físicos».

Bibliografía

Bernkopf, Michael: «A history of infinite matrices», *Archive for History of Exact Sciences*, 4, 1968, 308-358.

Cayley, Arthur: *The collected mathematical papers*, 13 vols., Cambridge University Press (1889-1897), Johnson Reprint Corp., 1963.

Feldman, Richard W., Jr.: (Seis artículos sobre matrices con diversos títulos). *The Mathematics Teacher*, 55, 1962, 482-484, 589-590, 657-659; 56, 1963, 37-38, 101-102, 163-164.

Frobenius, F. G.: *Gesammelte Abhandlungen*, 3 vols., Springer-Verlag, 1968.

Jacobi, C. G. J.: *Gesammelte Werke*, Georg Reimer, 1884, vol. 3.

MacDuffee, C. C.: *The Theory of Matrices*, Chelsea, 1946.

Muir, Thomas: *The theory of determinants in the historical order of development* (1906-1923), 4 vols., Dover (reimpresión), 1960.

— : List of writings on the theory of matrices, *Amer. Jour. of Math.*, 20, 1898, 225-228.

Sylvester, James Joseph: *The Collected Mathematical Papers*, 4 vols., Cambridge University Press, 1904-1912.

Weierstrass, Karl: *Mathematische Werke*, Mayer und Muller, 1895, vol. 2.

Morris Kline

El pensamiento matemático de la Antigüedad a nuestros días, III

Versión española de: Alejandro Garciadiego, Mariano Martínez
Coordinación y revisión: Jesús Hernández

Alianza
Editorial

INDICE

Capítulo 34
LA TEORIA DE NUMEROS EN EL SIGLO XIX

> Es cierto que Fourier pensaba que el objeto principal de las matemáticas era el uso público y la explicación de los fenómenos naturales; pero un filósofo como él debía saber que el único objetivo de la ciencia es el honor del espíritu humano y que bajo esta perspectiva un problema de [la teoría de] números es tan valioso como un problema sobre el sistema del mundo.
>
> C. G. J. JACOBI

1. Introducción

Hasta el siglo XIX, la teoría de números era una serie de resultados aislados, aunque muchas veces brillantes. Una nueva era se inició con el *Disquisitiones Arithmeticae (Disquisiciones Aritméticas)* [1], de Gauss, que compuso a la edad de veinte años. Esta gran obra había sido enviada a la Academia Francesa en 1800 pero la rechazaron y entonces Gauss la publicó él mismo. En este libro estableció la notación; sistematizó y extendió la teoría existente; clasificó los problemas que debían ser estudiados y los métodos conocidos para atacarlos e introdujo nuevos métodos. En el trabajo de Gauss sobre la teoría de números hay tres ideas principales: la teoría de congruencias, la introducción de los números algebraicos y la teoría de las formas como idea básica en el análisis diofántico. Este trabajo no solamente inició la teoría moderna de números, sino que determinó las direcciones de trabajo en la materia hasta el presente. Las *Disquisitiones* son difíciles de leer, pero Dirichlet las explicó.

[1] Publicado 1801 = *Werke*, 1.

Otro avance importante en el siglo XIX es la teoría analítica de números, que además de usar el álgebra se vale del análisis para tratar problemas concernientes a los enteros. Dirichlet y Riemann se erigieron como figuras en esta innovación.

2. La teoría de congruencias

A pesar de que la noción de congruencia no se originó con Gauss —ya que aparece en los trabajos de Euler, Lagrange y Legendre—, introdujo la notación en la primera sección del *Disquisitiones*, aplicándola a partir de ahí sistemáticamente. La idea básica es simple: el número 27 es congruente con el 3 módulo 4,

$$27 \equiv 3 \text{ módulo } 4,$$

ya que $27 - 3$ es exactamente divisible por 4. (La palabra módulo con frecuencia es abreviada como mód.) En general, si a, b y m son enteros:

$$a \equiv b \text{ módulo } m$$

si $a - b$ es (exactamente) divisible por m o si a y b tienen los mismos restos (o residuos), en la división por m. Entonces se dice que b es un residuo de a módulo m y a es un residuo de b módulo m. Como Gauss muestra, todos los residuos de a módulo m, para una a y m fijas, están dados por $a + km$, donde $k = 0, \pm 1, \pm 2, \ldots$.

Las congruencias con respecto al mismo módulo se tratan hasta cierto punto como ecuaciones. Algunas congruencias pueden ser añadidas, sustraídas y multiplicadas. Es posible además preguntarnos acerca de la solución de congruencias donde haya incógnitas. Así, ¿qué valores de x satisfacen

$$2x \equiv 25 \text{ módulo } 12?$$

Esta ecuación no tiene soluciones ya que $2x$ es par y $2x - 25$ es impar. De aquí que $2x - 25$ no puede ser múltiplo de 12. El teorema básico sobre congruencias polinomiales, que Gauss vuelve a probar

en la segunda sección, ya lo había establecido Lagrange.[2] Una congruencia del grado n,

$$Ax^n + Bx^{n-1} + ... + Mx + N \equiv 0 \text{ módulo } p$$

cuyo módulo es un número primo p que no divide a A no puede tener más de n raíces no congruentes.

En la tercera sección Gauss toma los restos de potencias. Proporciona una prueba en términos de congruencias del teorema menor de Fermat, que, enunciado en términos de potencias, se lee: si p es primo y a no es múltiplo de p entonces:

$$a^{p-1} \equiv 1 \text{ módulo } p.$$

El teorema se sigue de su estudio de congruencias de grado superior, a saber,

$$x^n \equiv a \text{ módulo } m$$

donde a y m son primos entre sí. Este estudio fue continuado por muchos autores después de Gauss.

La cuarta sección de las *Disquisitiones* trata de los residuos cuadráticos. Si p es primo y a no es múltiplo de p y si existe una x tal que $x^2 \equiv a$ mód. p., entonces a es residuo cuadrático de p; si no es así, a no es un residuo cuadrático de p. Después de demostrar algunos teoremas subordinados sobre residuos cuadráticos, Gauss proporcionó la primera demostración rigurosa de la ley de reciprocidad cuadrática (cap. 25, sec. 4). Euler, a su vez, había dado un enunciado completo muy parecido al de Gauss en un ensayo de sus *Opuscula Analytica (Opúsculos Analíticos)* de 1783 (cap. 25, sec. 4). Sin embargo, en el parágrafo 151 de las *Disquisitiones,* Gauss afirma que nadie ha presentado el teorema en una forma tan simple como él lo ha hecho. Se refiere a otros trabajos de Euler, incluyendo otro ensayo en el *Opuscula,* y al trabajo de Legendre de 1785. De estos ensayos Gauss dice, correctamente, que las pruebas eran incompletas.

Se supone que Gauss descubrió una prueba de la ley en 1796 cuando tenía 19 años. Aportó otra demostración en las *Disquisitiones* y más tarde publicó otras cuatro. Entre sus ensayos inéditos se

---

[2] *Hist. de l'Acad. de Berlin,* 24, 1768, 192 sgs., pub. 1770 = *Œuvres,* 2, 655-726.

encontraron otras dos. Gauss afirma que buscó muchas pruebas porque deseaba hallar una que pudiera usarse para establecer el teorema de reciprocidad bicuadrática (véase más adelante). La ley de reciprocidad bicuadrática, que Gauss llamó la joya de la aritmética, es un resultado básico sobre congruencias. Después de que Gauss diera sus pruebas, otras 50 más fueron proporcionadas por matemáticos posteriores.

Gauss también trató las congruencias de polinomios. Si A y B son dos polinomios en x con, digamos, coeficientes reales, entonces se sabe que es factible encontrar polinomios únicos Q y R tales que:

$$A = B \cdot Q + R,$$

donde el grado de R es menor que el grado de B. Se dice entonces que dos polinomios A_1 y A_2 son congruentes módulo un tercer polinomio P si tienen el mismo resto R de la división por P.

Cauchy se valió de esta idea [3] para definir los números complejos usando congruencias polinomiales. Así, si $f(x)$ es un polinomio con coeficientes reales, entonces con la división por $x^2 + 1$:

$$f(x) \equiv a + bx \bmod x^2 + 1$$

ya que el residuo es de grado menor que el divisor. En este caso, a y b son necesariamente reales en virtud del proceso de división. Si $g(x)$ es otro polinomio:

$$g(x) \equiv c + dx \bmod x^2 + 1.$$

Ahora, Cauchy señala que si A_1, A_2 y B son polinomios cualesquiera y si

$$A_1 = BQ_1 + R_1 \text{ y } A_2 = BQ_2 + R_2,$$

entonces,

$$A_1 + A_2 \equiv R_1 + R_2 \bmod B, \text{ y } A_1A_2 \equiv R_1R_2 \bmod B.$$

---

[3] *Exercises d'analyse et de physique mathématique*, 4, 1847, 84 sgs. = *Œuvres* (1), 10, 312-323 y (2), 14, 93-120.

Ahora podemos ver fácilmente que

$$f(x) + g(x) \equiv (a + c) + (b + d)x \bmod x^2 + 1$$

y ya que $x^2 \equiv -1$ mod. $x^2 + 1$, que

$$f(x)g(x) \equiv (ac - bd) + (ad + bc)x \bmod x^2 + 1.$$

Vemos así que los números $a + bx$ y $c + dx$ se combinan como números complejos; esto es, tienen las propiedades formales de los números complejos, tomando x el lugar de i. Cauchy también demostró que todo polinomio $g(x)$ no congruente con 0 módulo $x^2 + 1$ tiene inverso, ello significa que existe un polinomio $h(x)$ tal que $h(x)g(x)$ es congruente con 1 módulo $x^2 + 1$.

Cauchy introdujo i por x, siendo i para él una cantidad indeterminada real. Entonces demostró que para cualquier

$$f(i) = a_0 + a_1 i + a_2 i^2 + \ldots$$

se tiene

$$f(i) \equiv a_0 - a_2 + a_4 - \ldots + (a_1 - a_3 + a_5 - \ldots)i \bmod i^2 + 1.$$

De aquí que cualquier expresión con números complejos se comporta como una de las formas $c + di$ y todo el aparato necesario para trabajar con expresiones complejas está disponible. Para Cauchy, entonces, los polinomios en i, con su concepción de i, toman el lugar de los números complejos y es posible situar dentro de una clase todos aquellos polinomios teniendo el mismo residuo módulo $i^2 + 1$. Estas clases son los números complejos.

Es interesante que en 1847 Cauchy aún tuviera sus dudas acerca de $\sqrt{-1}$. Dice: «En la teoría de equivalencias algebraicas sustituida por la teoría de números imaginarios, la letra i cesa de representar al signo simbólico $\sqrt{-1}$, que repudiamos completamente y el cual podemos abandonar sin remordimiento ya que no se sabe lo que este supuesto signo significa ni qué sentido atribuirle. Por el contrario, representamos por la letra i una cantidad real pero indeterminada, y al sustituir el signo $\equiv$ por el signo $=$ transformamos lo que ha sido llamado una ecuación imaginaria en una equivalencia

algebraica relativa a la variable i y al divisor $i^2 + 1$. Ya que el divisor permanece igual en todas las fórmulas, se evita el escribirlo.»

En la segunda década del siglo, Gauss procedió a buscar las leyes de reciprocidad aplicables a las congruencias de grado superior. Estas leyes introducen de nuevo residuos de congruencias. Así, para la congruencia

$$x^4 \equiv q \bmod p$$

se define q como un residuo bicuadrático de p si existe un valor entero de x satisfaciendo la ecuación. Llegó a la ley de reciprocidad cuadrática (véase más abajo) y a una ley de reciprocidad cúbica. Gran parte de este trabajo apareció en los artículos de 1808 a 1817 y el propio teorema sobre residuos bicuadráticos aparece en ensayos de 1828 y 1832 [4].

Para obtener elegancia y sencillez en su teoría de residuos cúbicos y bicuadráticos, Gauss hizo uso de enteros complejos, esto es, números de la forma $a + bi$ con a y b enteros o cero. En el trabajo de Gauss sobre residuos bicuadráticos fue necesario considerar el caso donde el módulo p es primo de la forma $4n + 1$ y Gauss necesitaba los factores complejos en los cuales los números primos de la forma $4n + 1$ podían ser descompuestos. Para obtener éstos, Gauss se percató de que se debía ir más allá del dominio de los enteros ordinarios para introducir los enteros complejos. A pesar de que Euler y Lagrange habían introducido tales enteros en la teoría de números, fue Gauss quien les dio su importancia.

Mientras que en la teoría ordinaria de los enteros las unidades son $+1$ y -1, en la teoría de Gauss de enteros complejos las unidades son ± 1 y $\pm i$. A un entero complejo se le llama compuesto si es el producto de dos de tales enteros, donde ninguno es una unidad. Si tal descomposición no es posible, el entero es llamado primo. Así, $5 = (1 + 2i)(1 - 2i)$ es compuesto, mientras que 3 es un primo complejo.

Gauss demostró que los enteros complejos tienen esencialmente las mismas propiedades que los enteros ordinarios. Euclides había demostrado (cap. 4, sec. 7) que cada entero es descomponible uní-

[4] *Comm. Soc. Gott.*, 6, 1828 y 7, 1832 = *Werke*, 2, 65-92 y 93-148, también pp. 165-178.

vocamente en un producto de primos. Gauss demostró que esta descomposición única, a la que se llama frecuentemente el teorema fundamental de la aritmética, se mantiene también para enteros complejos siempre y cuando no pensemos en las cuatro unidades como factores diferentes. Esto es, si $a = bc = (ib)(-ic)$, las dos descomposiciones son la misma. Gauss demostró igualmente que el proceso de Euclides para encontrar el máximo común divisor de dos enteros es aplicable a los enteros complejos.

Muchos teoremas para primos ordinarios siguen siendo válidos para los primos complejos. Así, el teorema de Fermat se extiende en la forma: si p es un primo complejo $a + bi$, y k cualquier entero complejo no divisible por p, entonces

$$k^{Np-1} \equiv 1 \text{ módulo } p$$

donde Np es la norma $(a^2 + b^2)$ de p. También existe una ley de reciprocidad cuadrática para enteros complejos, establecida por Gauss en su ensayo de 1828.

En términos de los enteros complejos, Gauss fue capaz de enunciar con simplicidad la ley de reciprocidad bicuadrática. Un entero no par se define como uno no divisible entre $1 + i$. Un entero primariamente no par es un entero no par $a + bi$ tal que b es par y $a + b - 1$ es par. Así, -7 y $-5 + 2i$ son números primariamente no pares. La ley de reciprocidad para residuos bicuadráticos establece que si α y β son dos primos primariamente no pares y A y B son sus normas, entonces

$$\left(\frac{\alpha}{\beta}\right)_4 = (-1)^{(1/4)(A-1)(1/4)(B-1)}\left(\frac{\beta}{\alpha}\right)_4.$$

El símbolo $(\alpha/\beta)_4$ tiene el siguiente significado: si p es cualquier primo complejo y k es cualquier residuo bicuadrático no divisible por p, entonces $(k/p)_4$ es la potencia de i^e de i, que satisface la congruencia

$$k^{(Np-1)/4} \equiv 1 \text{ módulo } p$$

donde Np denota la norma de p. Esta ley es equivalente al enunciado: los caracteres bicuadráticos de dos números primos primarios no pares uno con respecto al otro son idénticos, esto es $(\alpha/\beta)_4 =$

pero si ninguno de los primos satisface la congruencia, entonces los dos caracteres bicuadráticos son opuestos, es decir $(\alpha/\beta)_4 = -(\beta/\alpha)_4$.

Gauss enunció este teorema de reciprocidad, pero no publicó su demostración. Esta fue proporcionada por Jacobi en sus clases en Königsberg en 1836-1837. Ferdinand Gotthold Eisenstein (1823-1852), un alumno de Gauss, publicó cinco demostraciones de la ley, de las cuales las dos primeras aparecieron en 1844.[5]

Para la reciprocidad cúbica, Gauss encontró que podía obtener una ley usando los «enteros» $a + b\varrho$, donde ϱ es una raíz de $x^2 + x + 1 = 0$ y a y b son dos enteros (racionales) ordinarios, pero Gauss no publicó este resultado, el cual se halló entre sus papeles después de su muerte. La ley de reciprocidad cúbica la establece por primera vez Jacobi[6] y es demostrada por él en sus clases en Königsberg. La primera demostración publicada se le debe a Eisenstein.[7] Al conocer esta demostración, Jacobi[8] reclamó que era precisamente la que él había dado en sus clases, pero Eisenstein negó indignadamente cualquier plagio.[9] También hay leyes de reciprocidad para congruencias de grado mayor que cuatro.

3. Números algebraicos

La teoría de enteros complejos es un paso en dirección de un dominio muy vasto, la teoría de los números algebraicos. Ni Euler ni Lagrange vieron las ricas posibilidades que abrió su trabajo con enteros complejos. Tampoco lo hizo Gauss.

La teoría surgió de los intentos por demostrar la aserción de Fermat acerca de $x^n + y^n = z^n$. Los casos $n = 3$, 4 y 5 ya han sido discutidos (cap. 25, sec. 4). Gauss intentó demostrar la aserción para $n = 7$, pero falló. Tal vez porque estaba a disgusto con su fracaso, dijo en una carta de 1816 a Heinrich W. M. Olbers (1758-1840): «Confieso, por supuesto, que el teorema de Fermat como una proposición aislada tiene poca importancia para mí, ya que multitud de tales proposiciones, que uno no puede demostrar o refutar, pueden ser fácilmente formuladas.» Este caso particular de $n = 7$ fue resuel-

[5] *Jour. für Math.*, 28, 1844, 53-67 y 223-245.
[6] *Jour. für Math.*, 2, 1827, 66-69 = *Werke*, 6, 233-237.
[7] *Jour. für Math.*, 27, 1844, 289-310.
[8] *Jour. für Math.*, 30, 1846, 166-182, p. 172 = *Werke*, 6, 254-274.
[9] *Jour. für Math.*, 35, 135-274 (p. 273).

to por Lamé en 1839[10], y Dirichlet estableció la aserción para $n = 14$.[11] Sin embargo, la proposición general quedaba sin probar.

Ernst Eduard Kummer (1810-1893), quien se desvió de la teología a las matemáticas, se impuso como tarea probarla. Se convirtió en alumno de Gauss y Dirichlet y más tarde trabajó como profesor en Breslau y Berlín. A pesar de que su trabajo principal versó sobre teoría de números, hizo bellos descubrimientos en geometría que tuvieron su origen en problemas ópticos; también realizó importantes contribuciones al estudio de la refracción de la luz en la atmósfera.

Kummer tomó $x^p + y^p$, donde p es primo, y lo factorizó en

$$(x + y)(x + \alpha y) \ldots (x + \alpha^{p-1}y),$$

donde α es una raíz imaginaria p-ésima de la unidad. Ello significa que α es una raíz de

$$\alpha^{p-1} + \alpha^{p-2} + \ldots + \alpha + 1 = 0. \quad (1)$$

Esto lo llevó a extender la teoría de enteros complejos de Gauss a los números algebraicos en tanto que éstos son introducidos por ecuaciones tales como (1), esto es, números de la forma

$$f(\alpha) = a_0 + a_1\alpha + \ldots + a_{p-2}\alpha^{p-2},$$

donde cada a_i es un entero (racional) ordinario. (Ya que α satisface (1), los términos en α^{p-1} pueden ser reemplazados por términos con menores potencias). Kummer llamó a los números $f(\alpha)$ enteros complejos.

En 1843, Kummer dio definiciones apropiadas de enteros, enteros primos, divisibilidad, y así sucesivamente (proporcionaremos las definiciones habituales dentro de un momento) y entonces cometió el error de suponer que la factorización única se mantiene en la clase de los números algebraicos que él había introducido. Señaló, mientras le transmitía su manuscrito a Dirichlet en 1843, que esta suposición era necesaria para demostrar el teorema de Fermat. Dirichlet le informó de que la factorización única es válida sólo para ciertos primos p. Incidentalmente, Cauchy y Lamé cometieron el mismo

[10] *Jour. de Math.*, 5, 1840, 195-211.
[11] *Jour. für Math.*, 9, 1832, 390-393 = *Werke*, 1, 189-194.

error de suponer la factorización única para los números algebraicos. En 1844, Kummer[12] reconoció lo correcto de la crítica de Dirichlet.

Para restaurar la factorización única, Kummer creó la teoría de números ideales en una serie de ensayos iniciados en 1844.[13] A fin de entender su idea consideremos el dominio de $a + b\sqrt{-5}$, donde a y b son enteros. En este dominio

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

y se demuestra fácilmente que los cuatro factores son enteros primos. Por tanto la descomposición única no tiene lugar. Introduzcamos, para este dominio, los números ideales $\alpha = \sqrt{2}$, $\beta_1 = (1 + \sqrt{-5})/\sqrt{2}$, $\beta_2 = (1 - \sqrt{-5})/\sqrt{2}$. Vemos que $6 = \beta_1\beta_2$. Entonces 6 está ahora expresado unívocamente como el producto de cuatro factores, todos números ideales en cuanto se refiere al dominio $a - b\sqrt{-5}$.[14] En términos de estos ideales y otros primos la factorización en el dominio es única (aparte de los factores formados por unidades). Con los números ideales es posible demostrar algunos de los resultados de la teoría de números ordinaria en todos los dominios que previamente carecían de factorización única.

Los números ideales de Kummer, aunque números ordinarios, no pertenecen a la clase de números algebraicos que él había introducido. Más aún, los números ideales no fueron definidos en ninguna forma general. En cuanto se refiere al teorema de Fermat, Kummer tuvo éxito al demostrar con sus números ideales que era cierto para algunos números primos. De los primeros cien enteros, sólo el 37, 59 y 67 no eran cubiertos por la demostración de Kummer. Entonces, en su ensayo de 1857,[15] extendió sus resultados a estos primos excepcionales. Tales resultados los amplió aún más Dimitri Mirimanoff (1861-1945), profesor de la universidad de Ginebra, quien perfeccionó los métodos de Kummer.[16] Mirimanoff demostró que el teorema de Fermat es cierto para todo n hasta 256 si x, y y z son primos con ese exponente n.

Mientras que Kummer trabajó con números algebraicos forma-

---

[12] *Jour. de Math.*, 12, 1847, 185-212.

[13] *Jour. für Math.*, 35, 1847, 319-326, 327-367.

[14] Con la introducción de los números ideales, 2 y 3 dejaron de ser descomponibles, ya que $2 = \alpha^2$ y $3 = \beta_1\beta_2$.

[15] *Abh. König. Akad. der Wiss. Berlin*, 1858, 41-74.

[16] *Jour. für Math.*, 128, 1905, 45-68.

dos a partir de las raíces de la unidad, Richard Dedekind (1831-1916), un alumno de Gauss —quien dedicó cincuenta años de su vida a enseñar en un instituto técnico de Alemania— estudió el problema de la factorización única de una manera enteramente nueva y fresca. Dedekind publicó sus resultados en el suplemento 10 a la segunda edición de la *Zahlentheorie (Teoría de Números)* de Dirichlet, que Dedekind editó. Difundió estos resultados en los suplementos de las tercera y cuarta ediciones del mismo libro.[17] Allí creó la teoría moderna de los números algebraicos.

La teoría de Dedekind de los números algebraicos es una generalización de los enteros complejos de Gauss y los números algebraicos de Kummer, pero la generalización está de alguna manera divorciada de los enteros complejos de Gauss. Un número r que es una raíz de

$$a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0, \qquad (2)$$

donde las a_i son enteros ordinarios (positivos o negativos), y que no es una raíz de una ecuación de grado menor que n es llamado un número algebraico de grado n. Si el coeficiente de la máxima potencia de x en (2) es 1, las soluciones son llamadas enteros algebraicos de grado n. La suma, diferencia y producto de enteros algebraicos son enteros algebraicos, y si un entero algebraico es un número racional, es un entero ordinario.

Debemos señalar que con las nuevas definiciones un entero algebraico puede contener fracciones ordinarias. Así $(-3 + \sqrt{-115})/2$ es un entero algebraico de segundo grado porque se trata de una raíz de $x^2 + 13x + 71 = 0$. Por otro lado, $(1 - \sqrt{-5})/2$ es un número algebraico de grado 2 pero no un entero algebraico, porque es una raíz de $2x^2 - 2x + 3 = 0$.

Dedekind introdujo después el concepto de cuerpo de números, que es una colección F de números reales o complejos tal que si α y β pertenecen a F, entonces también lo hacen $\alpha + \beta$, $\alpha - \beta$, $\alpha\beta$ y, si $\beta \neq 0$, α/β. Todo cuerpo de números contiene a los números racionales, ya que si α pertenece, entonces también α/α ó 1 y consecuentemente 1 + 1, 1 + 2, y de ahí en adelante. No es difícil demostrar que el conjunto de todos los números algebraicos forma un cuerpo.

[17] 4.ª ed., 1894 = *Werke*, 3, 2-222.

Si se empieza con el cuerpo de los números racionales y θ es un número algebraico de grado n, entonces el conjunto formado por la combinación de θ consigo mismo y los números racionales bajo las cuatro operaciones es también un cuerpo de grado n. Este cuerpo puede ser descrito alternativamente como el cuerpo mínimo conteniendo los números racionales y θ. También se le denomina un cuerpo de extensión de los números racionales. Tal cuerpo no consiste en todos los números algebraicos y es un campo de números algebraicos específico. La notación $R(\theta)$ es ahora común. Aunque fuera razonable esperar que los miembros de $R(\theta)$ son los cocientes $f(\theta)/g(\theta)$, donde $f(x)$ y $g(x)$ son polinomios cualesquiera con coeficientes racionales, se puede demostrar que si θ es de grado n, entonces cualquier miembro α de $R(\theta)$ se expresa en la forma

$$\alpha = a_0\theta^{n-1} + a_1\theta^{n-2} + \dots + a_{n-1},$$

donde las a_i son números racionales ordinarios. Más aún, existen enteros algebraicos $\theta_1, \theta_2, \dots, \theta_n$ de este cuerpo tales que los enteros algebraicos del cuerpo son de la forma

$$A_1\theta_1 + A_2\theta_2 + \dots + A_n\theta_n,$$

donde las A_i son enteros ordinarios positivos y negativos.

Un anillo, concepto introducido por Dedekind, es esencialmente cualquier colección de números tales que si α y β pertenecen, lo hacen asimismo $\alpha + \beta$, α-β, y $\alpha\beta$. El conjunto de todos los enteros algebraicos forma un anillo, al igual que lo hace el conjunto de todos los enteros algebraicos de cualquier cuerpo de números algebraicos específico.

Se dice que el entero algebraico α es divisible por el entero algebraico β si existe un entero algebraico γ tal que $\alpha = \beta\gamma$. Si γ es un entero algebraico que divide cualquier otro entero de un cuerpo de números algebraicos, entonces γ es llamada una unidad de ese cuerpo. Estas unidades, que incluyen al $+1$ y -1, son una generalización de las unidades $+1$ y -1 de la teoría de números ordinaria. El entero algebraico α es primo si no es cero o una unidad y si cualquier factorización de α en $\beta\gamma$, donde β y γ pertenecen al mismo cuerpo de números algebraicos, implica que β o γ es una unidad en ese cuerpo.

Vemos ahora hasta qué punto se mantiene el teorema fundamen-

tal de la aritmética. En el anillo de *todos* los enteros algebraicos no hay primos. Consideremos el anillo de los enteros en un cuerpo $R(\theta)$ específico de números algebraicos, digamos el cuerpo $a + b\sqrt{-5}$, donde a y b son dos números racionales ordinarios. En este cuerpo la factorización única no se mantiene. Por ejemplo,

$$21=3\cdot 7=(4+\sqrt{-5})(4-\sqrt{-5})=(1+2\sqrt{-5})(1-2\sqrt{-5}).$$

Cada uno de estos últimos cuatro factores es primo en el sentido de que no se expresa como un producto de la forma $(c + d\sqrt{-5})$ $(e + f\sqrt{-5})$ con c, d, e y f enteros.

Por otra parte, consideremos el cuerpo $a + b\sqrt{6}$, donde a y b son números racionales ordinarios. Si se aplican las cuatro operaciones algebraicas a estos números se obtienen números de la misma clase. Si a y b son restringidos a los enteros, se obtienen los enteros algebraicos (de grado 2) de este dominio. En este dominio podemos tomar como definición equivalente de unidad que el entero algebraico M es una unidad si $1/M$ es también un entero algebraico. Así 1, -1, $5 - 2\sqrt{6}$ y $5 + 2\sqrt{6}$ son unidades. Todo entero es divisible por cualquiera de las unidades. Además, un entero algebraico del dominio es primo si es divisible únicamente por él mismo y las unidades. Ahora,

$$6 = 2\cdot 3 = \sqrt{6}\cdot\sqrt{6}.$$

Parecería como si no existiera una descomposición única en primos. Pero los factores mostrados no son primos. De hecho

$$6 = 2\cdot 3 = \sqrt{6}\cdot\sqrt{6} = (2 + \sqrt{6})(-2 + \sqrt{6})(3 + \sqrt{6})(3 - \sqrt{6}).$$

Cada uno de los últimos cuatro factores es primo en el dominio y la descomposición única se cumple en el dominio.

En el anillo de los enteros de un cuerpo de números algebraicos la factorización específica de los enteros algebraicos en primos siempre es posible, pero la factorización *única* no siempre lo es. De hecho, para los dominios de la forma $a + b\sqrt{-D}$, donde D puede tener cualquier valor entero positivo no divisible por un cuadrado, el teorema de factorización única es válido sólo cuando $D = 1$, 2,

3, 7, 11, 19, 43, 67 y 163, al menos para D hasta 10^9.[18] Así, los propios números algebraicos no poseen la propiedad de la factorización única.

4. Los ideales de Dedekind

Después de generalizar la noción de número algebraico, Dedekind se dedicó a restaurar la factorización única en los cuerpos de números algebraicos mediante un esquema bastante diferente al de Kummer. En lugar de números ideales, Dedekind introdujo las clases de números algebraicos que llamó ideales, en honor a los números ideales de Kummer.

Antes de definir los ideales de Dedekind, señalemos cuál es el pensamiento subyacente. Considérense los enteros ordinarios. En lugar del entero 2, Dedekind toma la clase de los enteros $2m$, donde m es cualquier entero. Esta clase consiste de todos los enteros divisibles por 2. De la misma manera, el 3 es reemplazado por la clase de todos los enteros $3n$ divisibles por 3. El producto 6 se convierte en la colección de todos los números $6p$, donde p es cualquier entero. Entonces el producto $2 \cdot 3 = 6$ es reemplazado por la afirmación de que la clase $2m$ «multiplicada» por la clase $3n$ es igual a la clase $6p$. Más aún, la clase $2m$ es un factor de la clase $6p$, a pesar del hecho de que la clase anterior contiene a la última. Estas clases son ejemplos en el anillo de los enteros ordinarios de lo que Dedekind llama ideales. Para seguir el trabajo de Dedekind, uno debe acostumbrarse a pensar en términos de clases de números.

Con mayor generalidad, Dedekind definió los ideales como sigue: sea K un cuerpo de números algebraicos específico. Un conjunto de enteros A de K se dice que forma un ideal si cuando α y β son dos enteros cualesquiera en el conjunto, los enteros $\mu\alpha + \nu\beta$, donde μ y ν son otros dos enteros cualesquiera en K, también pertenecen al conjunto. Alternativamente, un ideal A se dice que es generado por los enteros algebraicos $\alpha_1, \alpha_2, \ldots, \alpha_\nu$ de K si A consiste

[18] H. M. Stark ha demostrado que los valores anteriores de D son los únicos posibles. Véase su «On the problem of unique factorization in complex quadratic fields», *Proceedings of Symposia in Pure Mathematics*, XII, 41-56, Amer. Math. Soc., 1969.

en todas las sumas

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + ... + \lambda_n\alpha_n,$$

donde los λ_r son enteros cualesquiera de un cuerpo K. Este ideal es denotado mediante $(\alpha_1, \alpha_2, ..., \alpha_v)$. El ideal cero consiste en el número 0 únicamente y por consiguiente es denotado por (0). El ideal unidad es el generado por el número 1, esto es, (1). Un ideal A es llamado principal si es generado por un entero único α, de tal forma que (α) consta de todos los enteros algebraicos divisibles por α. En el anillo de los enteros algebraicos todo ideal es un ideal principal.

Un ejemplo de un ideal en el campo $a + b\sqrt{5}$ de los números algebraicos, donde a y b son números racionales ordinarios, es el ideal generado por los enteros 2 y $1 + \sqrt{-5}$. Este ideal consiste en todos los enteros de la forma $2\mu + (1 + \sqrt{-5})\, \nu$, donde μ y ν son enteros arbitrarios del cuerpo. El ideal también es un ideal principal, ya que está generado sólo por el número 2 en virtud del hecho de que $(1 + \sqrt{-5})\, 2$ debe pertenecer también al ideal generado por 2.

Dos ideales $(\alpha_1, \alpha_2, ..., \alpha_p)$ y $(\beta_1, ... \beta_q)$ son iguales si cada miembro del primer ideal es un miembro del segundo y conversamente. Para atacar el problema de la factorización debemos primero considerar el producto de dos ideales. El producto del ideal $A = (\alpha_1, ..., \alpha_s)$ y el ideal $B = (\beta_1, ..., \beta_t)$ de K está definido como el ideal

$$AB = (\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, ..., \alpha_i\beta_j, ..., \alpha_s\beta_t).$$

Es casi evidente que este producto es conmutativo y asociativo. Con la definición podemos decir que A divide a B si existe un ideal C tal que $B = AC$. Se escribe A/B y A es llamado un factor de B. Como ya se había sugerido mediante nuestro ejemplo de los enteros ordinarios, los elementos de B están *incluidos* en los elementos de A y la divisibilidad ordinaria es reemplazada por la inclusión de clases.

Los ideales que son los análogos de los números primos ordinarios son llamados ideales primos. Tal tipo de ideal P es definido como uno que no tiene factores que no sean él mismo y el ideal (1), de tal forma que P no está contenido en ningún otro ideal de K. Por esta razón un ideal primo es también llamado maximal. Todas estas definiciones y teoremas condujeron a los teoremas básicos para ideales de un cuerpo de números algebraicos K. Cualquier ideal es divisible únicamente por un número finito de ideales y si un ideal

primo divide el producto AB de dos ideales (de la misma clase de números) divide a A o a B. Finalmente, el teorema fundamental en la teoría de ideales es que todo ideal puede ser factorizado unívocamente en ideales primos.

En nuestros primeros ejemplos de cuerpos de números algebraicos de la forma $a + b\sqrt{D}$, siendo entero D, encontramos que algunos permitían la factorización única de los enteros algebraicos de aquellos cuerpos y otros no. La respuesta a la pregunta de cuáles lo permiten y cuáles no está dada por el teorema de que la factorización de los enteros de un cuerpo K de números algebraicos en primos es única si y sólo si todos los ideales de K son principales.

A partir de estos ejemplos del trabajo de Dedekind podemos ver que su teoría de los ideales es por supuesto una generalización de los enteros ordinarios. En particular, suministra los conceptos y propiedades en el dominio de los números algebraicos que permiten establecer la factorización única.

Leopold Kronecker (1823-1891), fue el estudiante favorito de Kummer y le sucedió como profesor en la universidad de Berlín; también se dedicó al estudio de los números algebraicos y lo desarrolló siguiendo líneas similares a las de Dedekind. La tesis doctoral de Kronecker «Sobre unidades complejas», escrita en 1845, aunque no publicada sino mucho más tarde,[19] fue su primer trabajo en el tema. La tesis trata de las unidades que pueden existir en los cuerpos de números algebraicos creados por Gauss.

Kronecker creó otra teoría de cuerpos (dominios de racionalidad).[20] Su concepto de cuerpo es mucho más general que el de Dedekind, ya que consideró campos de funciones racionales en cualquier número de variables (indeterminadas). Específicamente, Kronecker introdujo (1881) la noción de una indeterminada adjunta a un cuerpo siendo únicamente la indeterminada una nueva cantidad abstracta. Esta idea de extender un cuerpo añadiendo una indeterminada la convirtió en la piedra angular de su teoría de números algebraicos. Aquí utilizó el conocimiento ya construido desde Liouville, Cantor y otros, sobre la distinción entre los números algebraicos y los trascendentes. En particular, observó que si x es trascendente sobre un

---

[19] *Jour. für Math.*, 93, 1882, 1-52 = *Werke*, 1, 5-71.

[20] «Grundzüge einer arithmetischen Theorie der algebraischen Grössen», *Jour. für Math.*, 92, 1882, 1-122 = *Werke*, 2, 237-387; también publicado por G. Reimer, 1882.

cuerpo K (x es una indeterminada) entonces el campo $K(x)$ obtenido al añadir la indeterminada x a K, esto es, el cuerpo más pequeño conteniendo K y x, es isomorfo al cuerpo $K[x]$ de funciones racionales de una variable con coeficientes en K.[21] Subrayó que la indeterminada era únicamente un elemento de un álgebra y no una variable en el sentido del análisis.[22] Más adelante, en 1887,[23] demostró que a cada número primo ordinario p le corresponde dentro del anillo $A(x)$ de polinomios con coeficientes racionales un polinomio primo $p(x)$ que es irreducible en el campo racional Q. Considerando dos polinomios como iguales si son congruentes módulo un polinomio primo $p(x)$ dado, el anillo de los polinomios en $Q(x)$ se convierte en un cuerpo de clases de restos poseyendo las mismas propiedades algebraicas que el cuerpo de los números algebraicos $K(\delta)$ obtenido del cuerpo K al añadir una raíz δ de $p(x) = 0$. Aquí utilizó la idea que Cauchy ya había empleado para introducir los números imaginarios al utilizar polinomios congruentes módulo $x^2 + 1$. En este mismo trabajo demostró que la teoría de los números algebraicos es independiente del teorema fundamental del álgebra y de la teoría del sistema completo de los números reales.

En su teoría de cuerpos (en los «Grundzüge»), cuyos elementos están formados empezando con un campo K y adjuntando indeterminadas $x_1, x_2, ..., x_n$, Kronecker introdujo la noción de un sistema modular que jugó el papel de los ideales en la teoría de Dedekind. Para Kronecker, un sistema modular es el conjunto M de aquellos polinomios en n variables $x_1, x_2, ... x_n$, tales que si P_1 y P_2 pertenecen al conjunto, también lo hace $P_1 + P_2$, y si P pertenece, también lo hace QP, donde Q es cualquier polinomio en $x_1, x_2, ..., x_n$.

Una base del sistema modular M es cualquier conjunto de polinomios $B_1, B_2...$ de M tal que cada polinomio de M es expresable en la forma

$$R_1B_1 + R_2B_2 + ...,$$

donde $R_1, R_2...$ son constantes o polinomios (no perteneciendo necesariamente a M). La teoría de la divisibilidad en los cuerpos gene-

[21] *Werke*, 2, 253.
[22] *Werke*, 2, 339.
[23] *Jour. für Math.*, 100, 1887, 490-510 = *Werke*, 3, 211-240.

rales de Kronecker fue definida en términos de sistemas modulares, un poco como lo había hecho Dedekind con los ideales.

El trabajo sobre la teoría algebraica de números llegó a su clímax en el siglo XIX con el famoso artículo de Hilbert sobre números algebraicos.[24] Este trabajo es esencialmente una descripción de lo que se había hecho durante el siglo. Sin embargo, Hilbert retrabajó toda esta teoría anterior y proporcionó nuevos, elegantes y poderosos métodos para asegurar sus resultados. Empezó a crear nuevas ideas en la teoría algebraica de números alrededor de 1892 aproximadamente, y una de las nuevas creaciones sobre los cuerpos de números galoisianos fue incorporada también en el trabajo. Posteriormente, Hilbert y muchos otros autores extendieron ampliamente la teoría de números algebraicos. Sin embargo, estos últimos desarrollos, relativos a cuerpos galoisianos, cuerpos de números abelianos relativos y cuerpos de clases, cada uno de ellos estimulando una inmensa cantidad de trabajo en el siglo XX, son de interés principalmente para los especialistas.

La teoría algebraica de números, que originalmente era un esquema para investigar las soluciones de los problemas de la vieja teoría de números, se ha convertido en un fin en sí misma. Ha llegado a ocupar una posición entre la teoría de números y el álgebra abstracta, y ahora la teoría de números y el álgebra superior moderna se mezclan en la teoría algebraica de números. Por supuesto, la teoría algebraica de números también ha producido nuevos teoremas en la teoría ordinaria de números.

5. La teoría de las formas

Otra clase de problemas en la teoría de números es la representación de los enteros por medio de formas. La expresión

$$ax^2 + 2bxy + cy^2, \qquad (3)$$

donde a, b y c son enteros, es una forma binaria ya que hay dos variables implicadas y es una forma cuadrática porque es de segundo grado. Se dice que un número M está representado por la forma si

---

[24] «Die Theorie der algebraischen Zahlkörper», *Jahres. der Deut. Math.-Ver.*, 4, 1897, 175-546 = *Ges. Abh.*, 1, 63-363.

para valores enteros concretos de a, b, c, x e y, la expresión anterior es igual a M. Uno de los problemas es encontrar el conjunto de números M que son representables por una forma dada o clase de formas. El problema inverso, dado M y dados a, b y c, o alguna clase de a, b y c, encontrar los valores de x y y que representa M, es igualmente importante. El último problema pertenece al análisis diofántico y el anterior puede ser igualmente considerado parte de la misma materia.

Euler había obtenido algunos resultados particulares sobre estos problemas. Sin embargo, Lagrange hizo el descubrimiento clave de que si un número es representable por una forma, es también representable por muchas otras formas, que llamó equivalentes. Las últimas podían ser obtenidas a partir de la forma original mediante un cambio de variables

$$x = \alpha x' + \beta y', \quad y = \gamma x' + \delta y' \tag{4}$$

donde las α, β, γ y δ son enteros y $\alpha\delta - \beta\gamma = 1$.[25] En particular, Lagrange mostró que para un discriminante dado (Gauss usó la palabra determinante) $b^2 - ac$ existe un número finito de formas tal que cada forma con ese discriminante es equivalente a una de ese número finito. De esta manera, todas las formas con un discriminante dado pueden ser divididas en clases, consistiendo cada clase en formas equivalentes a un miembro de esa clase. Este resultado y otros obtenidos inductivamente por Legendre atrajeron la atención de Gauss. En un paso audaz, Gauss extrajo del trabajo de Lagrange la noción de equivalencia de formas y se concentró en ella. La quinta sección de sus *Disquisitiones*, con mucho la sección más larga, está dedicada a este tema.

Gauss sistematizó y extendió la teoría de las formas. Primero definió la equivalencia de formas. Sea

$$F = ax^2 + 2bxy + cy^2$$

transformada por medio de (4) en la forma

$$F' = a'x^{2\prime} + 2b'x'y' + c'y'^2.$$

[25] *Nouv. Mém. de l'Acad. de Berlin*, 1773, 263-312; y 1775, 323 sgs. = *Œuvres*, 3, 693-795.

Entonces

$$b'^2 - a'c' = (b^2 - ac)(\alpha\delta - \beta\gamma)^2.$$

Si ahora $(\alpha\delta - \beta\gamma)^2 = 1$, los discriminantes de las dos formas son iguales. Entonces el inverso de la transformación (4) también contendrá coeficientes enteros (por la regla de Cramer) y transformará F' en F. Se dice que F y F' son equivalentes. Si $\alpha\delta - \beta\gamma = 1$, se dice que F y F' son propiamente equivalentes, y si $\alpha\delta - \beta\gamma = -1$, entonces se dice que F y F' son impropiamente equivalentes.

Gauss demostró un buen número de teoremas sobre la equivalencia de formas. Por ejemplo, si F es equivalente a F' y F' a F'', entonces F es equivalente a F''. Si F es equivalente a F', cualquier número M representable por F es representable por F' y de tantas maneras por una como por la otra. Más adelante demuestra, si F y F' son equivalentes, cómo encontrar todas las transformaciones de F en F'. También encuentra todas las representaciones de un número dado M por la forma F, siempre que los valores de x y y sean primos entre sí.

Por definición, dos formas equivalentes tienen el mismo valor para su discriminante $D = b^2 - ac$. Sin embargo, dos formas con el mismo discriminante no son necesariamente equivalentes. Gauss demuestra que todas las formas con un D dado pueden ser divididas en clases; los miembros de cualquier clase son propiamente equivalentes entre sí. Aunque el número de formas con un D dado es infinito, el número de clases para una D dada es finito. En cada clase puede ser tomada una forma como representativa y Gauss proporciona criterios para la elección de la representación más simple. La forma más simple de todas aquellas con determinante D tiene $a = 1$, $b = 0$ y $c = -D$. A esto llama Gauss la forma principal y la clase a la que pertenece es llamada clase principal.

Gauss, entonces, ataca la composición (producto) de formas. Si la forma

$$F = AX^2 + 2BXY + CY^2$$

se transforma en el producto de dos formas

$$f = ax^2 + 2bxy + cy^2 \text{ y } f' = a'x'^2 + 2b'x'y' + c'y'^2$$

por la sustitución

$$X = p_1xx' + p_2xy' + p_3x'y + p_4yy'$$
$$Y = q_1xx' + q_2xy' + q_3x'y + q_4yy'$$

entonces se dice que F es transformable en ff'. Si además los seis números

$$p_1q_2-q_1p_2,\ p_1q_3-q_1p_3,\ p_1q_4-q_1p_4,\ p_2q_3-q_2p_3,\ p_2q_4-q_2p_4,\ p_3q_4-q_3p_4$$

no tienen un divisor común, entonces se dice que F es compuesta de las formas f y f'.

Más adelante, Gauss demostró un teorema esencial: si f y g pertenecen a la misma clase y f' y g' pertenecen a la misma clase, entonces la forma compuesta de f y f' pertenecerá a la misma clase que la forma compuesta de g y g'. Entonces se puede hablar de una clase de formas compuesta de dos (o más) clases dadas. En esta composición de clases, la clase principal actúa como una clase unidad; esto es, si la clase K es compuesta con la clase principal, la clase resultante será K.

Gauss pasa ahora al tratamiento de las formas cuadráticas ternarias

$$Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2,$$

donde los coeficientes son enteros y realiza un estudio muy parecido al que había llevado a cabo para formas binarias. La meta, como en el caso de las formas binarias, es la representación de los enteros. Gauss no desarrolló la teoría de las formas ternarias.

El objetivo del trabajo entero de la teoría de formas fue, como se mencionó con anterioridad, obtener teoremas en teoría de números. En el curso de su tratamiento de las formas, Gauss muestra cómo la teoría puede ser usada para demostrar un gran número de teoremas sobre los enteros incluyendo muchos que habían sido previamente demostrados por autores tales como Euler y Lagrange. Así, Gauss demuestra que cualquier número primo de la forma $3n + 1$ puede ser representado como una suma de cuadrados de una manera única. Cualquier número primo de la forma $8n + 1$ o $8n + 3$ puede ser representado en la forma $x^2 + 2y^2$ (para enteros positivos x e y) en una y sólo una manera. Muestra cómo encontrar todas las repre-

sentaciones de un número dado M mediante la forma dada $ax^2 + 2bxy + cy^2$, siempre que el discriminante D sea un número positivo que no es un cuadrado. Además, si F es una forma primitiva (los valores de a, b y c son primos entre sí) con el discriminante D y si p es un número primo que divide a D, entonces los números no divisibles por p que pueden ser representados por F coinciden en que todos ellos o son residuos cuadráticos de p o no residuos de p.

Entre los resultados que Gauss dedujo de su trabajo sobre formas cuadráticas ternarias está la primera prueba del teorema de que cada número puede ser representado como suma de tres números triangulares. Estos, recordamos, son los números

$$1, 3, 6, 10, 15, \ldots, \frac{n^2 + n}{2}, \ldots$$

También volvió a demostrar el teorema ya probado por Lagrange de que cualquier entero positivo puede ser expresado como la suma de cuatro cuadrados. A propósito del primer resultado, es importante notar que Cauchy leyó un ensayo ante la Academia de París en 1815 donde establecía el resultado general (afirmado primero por Fermat) de que cada entero es la suma de k, o menos, números k-gonales.[26] (El número k-gonal general es $n + (n^2 - n)(k - 2)/2$).

La teoría algebraica de las formas cuadráticas binarias y ternarias —como la presentó Gauss— tiene una análogía geométrica interesante que inició el propio Gauss. En una reseña, que apareció en el *Göttingische Gelehrte Anzaigen* de 1830,[27] de un libro sobre formas cuadráticas ternarias, cuyo autor fuera Ludwig August Seeber (1793-1855), Gauss bosquejó la representación geométrica de las formas y de las clases de formas.[28] Este trabajo es el inicio de una rama llamada la teoría geométrica de números, que ganó primero importancia cuando Hermann Minkowski (1864-1909), quien trabajó como profesor en varias universidades, publicó su *Geometrie der Zahlen* (*Geometría de los Números*, 1896).

El campo de las formas adquirió importancia en la teoría de

---

[26] *Mém. de l'Acad. des Sci., Paris* (1), 14, 1813-1815, 177-220 = *Œuvres* (2), 6, 320-353.

[27] *Werke*, 2, 188-196.

[28] Felix Klein en su *Entwicklung* (véase la bibliografía al final de este capítulo), pp. 35-39, desarrolla el bosquejo de Gauss.

números del siglo XIX. Trabajo posterior sobre formas cuadráticas binarias y ternarias y sobre formas con más variables y de grado superior,[29] fue realizado por multitud de autores.

6. Teoría analítica de números

Uno de los principales avances en la teoría de números es la presentación de métodos analíticos, y de resultados analíticos, para expresar y demostrar hechos acerca de los enteros. De hecho, Euler había usado el análisis en la teoría de números (véase más adelante) y Jacobi utilizó funciones elípticas para obtener resultados en la teoría de congruencias y en la teoría de formas.[30] Sin embargo, los usos de Euler del análisis en la teoría de números fueron menores y los resultados de Jacobi, casi productos accidentales de su trabajo en análisis.

El primer uso deliberado y profundo del análisis para atacar lo que parecía ser un claro problema de álgebra fue realizado por Peter Gustav Lejeune-Dirichlet (1805-1859), alumno de Gauss y Jacobi, profesor en Breslau y Berlín, y más adelante sucesor de Gauss en Göttingen. El gran libro de Dirichlet, *Vorlesungen über Zahlentheorie (Lecturas sobre la teoría de números)*,[31] explicó las *Disquisitiones* de Gauss y dio sus propias contribuciones.

El problema que llevó a Dirichlet a emplear el análisis fue demostrar que cada sucesión aritmética

$$a, a + b, a + 2b, a + 3b, \ldots, a + nb, \ldots,$$

donde a y b son primos entre sí, contiene un número infinito de primos. Euler[32] y Legendre[33] hicieron esta conjetura y en 1808 Legendre[34] proporcionó una demostración que era errónea. En 1837, Dirichlet[35] dio una demostración correcta. Este resultado generaliza

[29] Para más detalles, véanse los trabajos de Smith y Dickson, incluidos en la bibliografía.

[30] *Jour. für Math.*, 37, 1848, 61-94 y 221-254 = *Werke*, 2, 219-288.

[31] Publicado en 1863; Dedekind suplementó extensamente las segunda, tercera y cuarta ediciones de 1871, 1879 y 1894.

[32] *Opuscula Analytica*, 2, 1783.

[33] *Mém. de l'Acad. des Sci., Paris*, 1785, 465-559, pub. 1788.

[34] *Théorie des nombres*, 2.ª ed., p. 404.

[35] *Abh. König. Akad. der Wiss., Berlin*, 1837, 45-81 y 108-110 = *Werke*, 1, 307-342.

el teorema de Euclides sobre la infinitud de primos en la sucesión 1, 2, 3, ... La demostración analítica de Dirichlet era larga y complicada. Específicamente, usaba lo que ahora son llamadas las series de Dirichlet, $\sum_{n=1}^{\infty} a_n n^{-z}$, donde las a_n y z son complejos. Dirichlet también demostró que la suma de los recíprocos de los primos en la sucesión $a + nb$ diverge. Esto extiende un resultado de Euler sobre los primos usuales (véase más abajo). En 1841 [36] Dirichlet demostró un teorema sobre los primos en progresiones de números complejos $a + bi$.

El problema principal que implicaba la introducción del análisis concernía a la función $\pi(x)$, que representa el número de primos que no exceden a x. Así $\pi(8)$ es 4 ya que 2, 3, 5 y 7 son primos, y $\pi(11)$ es 5. Cuando x se incrementa, los primos adicionales se hacen escasos y el problema era: ¿cuál es la expresión analítica para $\pi(x)$? Legendre, quien había demostrado que ninguna expresión racional puede servir, en algún momento perdió la esperanza de que pudiera ser encontrada cualquier expresión. Entonces Euler, Legendre, Gauss y otros supusieron que

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1. \tag{5}$$

Gauss utilizó tablas de primos (de hecho estudió todos los primos hasta 3.000.000) para hacer conjeturas acerca de $\pi(x)$ e infirió [37] que $\pi(x)$ difiere poco de $\int_2^x dt/\log t$. También sabía que

$$\lim_{x\to\infty} \frac{\int_2^x dt/\log t}{x/\log x} = 1.$$

En 1848, Pafnuti L. Tchebycheff (1821-1894), profesor de la universidad de Petrogrado, abordó la cuestión del número de números primos menor o igual a x y dio un paso importante hacia adelante

[36] *Abh. König. Akad. der Wiss., Berlin,* 1841, 141-161 = *Werke,* 2, 509-532.
[37] *Werke,* 2, 444-447.

en este viejo problema. En un ensayo clave, «Sur les nombres premiers» («Sobre los números primos»),[38] Tchebycheff demostró que

$$A_1 < \frac{\pi(x)}{x/\log x} < A_2,$$

donde 0,922 $<A_1$ <1 y 1$<A_2$ <1.105, pero no demostró que la función tiende a un límite. Esta desigualdad fue mejorada por muchos matemáticos, incluyendo a Sylvester, quien, entre otros, dudaba en 1881 de que la función tuviera un límite. En su trabajo, Tchebycheff usó lo que ahora conocemos como la función zeta de Riemann,

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$$

aunque la utilizó para valores reales de z. (Esta serie es un caso especial de la serie de Dirichlet.) Incidentalmente, también demostró en el mismo ensayo que para $n > 3$ siempre existe al menos un primo entre n y $2n - 2$.

La función zeta para z real aparece en el trabajo de Euler [39], donde introdujo

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \prod_{n=1}^{\infty} \left(1 - \frac{1}{p_n^s}\right)^{-1}.$$

Aquí los p_n son números primos. Euler utilizó la función para demostrar que la suma de los recíprocos de los números primos diverge. Para s entero par positivo, Euler conocía el valor de $\zeta(s)$ (cap. 20, sec. 4). Entonces, en un artículo leído en 1749 [40], Euler aseguró que para s real

$$\zeta(1 - s) = 2(2\pi)^{-s} \cos \frac{\pi s}{2}\, \Gamma(s)\zeta(s).$$

---

[38] *Mém. Acad. Sci. St. Peters.*, 7, 1854, 15-33; también en *Jour. de Math.* (1), 17, 1852, 366-390 = *Œuvres*, 1, 51-70.

[39] *Comm. Acad. Sci. Petrop.*, 9, 1737, 160-188, pub. 1744 = *Opera*, (1), 14, 216-244.

[40] *Hist. de l'Acad. de Berlin*, 17, 1761, 83-106, pub. 1768 = *Opera* (1), 15, 70-90.

Verificó la ecuación hasta el punto de que, decía, no había duda acerca de ella. Esta relación fue establecida por Riemann en el artículo de 1859, al que nos habíamos referido con anterioridad. Riemann, usando la función zeta para z complejo, intentó demostrar el teorema de los números primos, esto es, (5) antes mencionado.[41] Señaló que, para llevar más lejos la investigación, se tenían que conocer los zeros complejos de $\zeta(z)$. De hecho, $\zeta(z)$ no converge para $x \leqslant 1$ cuando $z = x + iy$, pero los valores de ζ en la mitad del plano $x \leqslant 1$ están definidos por continuación analítica. Expresó la hipótesis de que todos los zeros en la banda $0 \leqslant x \leqslant 1$ están sobre la línea $x = 1/2$. Esta hipótesis sigue sin ser demostrada.[42]

En 1896, Hadamard,[43] aplicando la teoría de funciones enteras (de una variable compleja), que investigó con el propósito de demostrar el teorema de los números primos y probando el hecho crucial que $\zeta(z) \neq 0$ para $x = 1$, estuvo finalmente en condiciones de demostrar el teorema de los números primos. Charles-Jean de la Vallée Poussin (1866-1962) obtuvo el resultado para la función zeta y demostró el teorema de los números primos en el mismo momento.[44] Este teorema es central en la teoría analítica de números.

Bibliografía

Bachmann, P.: «Über Gauss' zahlentheoretische Arbeiten». *Nachrichten Köning. Ges. der Wiss. zu Gött.*, 1911, 455-508; también en Gauss: *Werke*, 10, 1-69.

Bell, Eric T.: *The development of mathematics*, 2.ª ed., McGraw-Hill, 1945, caps. 9-10.

Carmichael, Robert D.: «Some recent researches in the theory of numbers». *Amer. Math. Monthly*, 39, 1932, 139-160.

Dedekind, Richard: *Über die Theorie der ganzen algebraischen Zahlen* (reimpreso del undécimo suplemento del *Zahlentheorie* de Dirichlet), F. Vieweg und Sohn, 1964.

— *Gesammelte mathematische Werke*, 3 vols., F. Vieweg und Sohn, 1930-1932, Chelsea (reimpresión), 1968.

---

[41] *Monatsber. Berliner Akad.*, 1859, 671-680 = *Werke*, 145-155.

[42] En 1914 Godfrey H. Hardy demostró (*Comp. Rend.*, 158, 1914, 1012-1014 = *Coll. Papers*, 2, 6-9) que una infinidad de ceros de $\zeta(z)$ está sobre la línea $x = 1/2$.

[43] *Bull. Soc. Math. de France*, 14, 1896, 199-220 = *Œuvres*, 1, 189-210.

[44] *Ann. Soc. Sci. Bruxelles* (1), 20 parte II, 1896, 183-256, 281-397.

— «Sur la théorie des nombres entiers algébriques», *Bull. des Sci. Math.*, (1), 11, 1876, 278-288; (2), 1, 1877, 17-41, 69-92, 144-164, 207-248 = *Ges. math. Werke*, 3, 263-296.
Dickson, Leonard E.: *History of the theory of numbers*, 3 vols., Chelsea (reimpresión), 1951.
— *Studies in the theory of numbers* (1930), Chelsea (reimpresión), 1962.
— y cols.: *Algebraic numbers, Report of Committee on algebraic numbers*, National Research Council, 1923 y 1928; Chelsea (reimpresión), 1967.
Dirichlet, P. G. L.: *Werke* (1889-1897); Chelsea (reimpresión), 1969, 2 vols.
Dirichlet, P. G. L., y Dedekind, R.: *Vorlesungen über Zahlentheorie*, 4.ª ed., 1894 (contiene el suplemento de Dedekind); Chelsea (reimpresión), 1968.
Gauss, C. F.: *Disquisitiones Arithmeticae*, traducción al inglés A. A. Clarke, Yale University Press, 1965.
Hasse, H.: «Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper», *Jahres. der Deut. Math.-Verein*, 35, 1926, 1-55 y 36, 1927, 233-311.
Hilbert, David: «Die Theorie der algebraischen Zahlkörper», *Jahres. der Deut. Math.-Verein*, 4, 1897, 175-546 = *Gesammelte Abhandlungen*, 1, 63-363.
Klein, Felix: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea (reimpresión), 1950, vol. 1.
Kronecker, Leopold: *Werke*, 5 vols. (1895-1931), Chelsea (reimpresión), 1968. Véase especialmente vol. 2, pp. 1-10 sobre la ley de reciprocidad cuadrática.
— *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*, G. Reimer, 1882 = *Jour. fur Math.*, 92, 1881/1882, 1-122 = *Werke*, 2, 237-388.
Landau, Edmund: *Handbuch der Lehre von der Verteilung der Primzahlen*, B. G. Teubner, 1909, vol. 1, págs. 1-55.
Mordell, L. J.: «An introductory account of the arithmetical theory of algebraic numbers and its recent development». *Amer. Math. Soc. Bull.*, 29, 1923, 445-463.
Reichardt, Hans (ed.): *C. F. Gauss, Leben und Werk*, Haude und Spenersche Verlagsbuchhandlund, 1960, pp. 38-91; también B. G. Teubner, 1957.
Scott, J. F.: *A history of mathematics*, Taylor and Francis, 1958, cap. 15.
Smith, David, E.: *A source book in mathematics*, Dover (reimpresión), 1959, vol. 1, 107-148.
Smith, H. J. S.: *Collected mathematical papers*, 2 vols. (1890-1894), Chelsea (reimpresión), 1965. El vol. 1 contiene el *Report on the theory of numbers* de Smith, que también ha sido publicado por separado en Chelsea, 1965.
Vandiver, H. S.: «Fermat's last theorem», *Amer. Math. Monthly*, 53, 1946, 555-578.

Capítulo 35
EL RESURGIMIENTO DE LA GEOMETRIA PROYECTIVA

Las doctrinas de la geometría pura, frecuentemente, y en muchas cuestiones, proporcionan una manera simple y natural de penetrar en los orígenes de las verdades, para aclarar la misteriosa cadena que las une, y para hacerlas conocer individual, luminosa y completamente.

MICHAEL CHASLES

1. El renovado interés por la geometría

Por más de 100 años después de la introducción de la geometría analítica por Descartes y Fermat, los métodos algebraicos y analíticos dominaron la geometría, hasta la casi exclusión de los métodos sintéticos. Durante este período algunos autores, por ejemplo los matemáticos ingleses que persistían en intentar fundar el cálculo rigurosamente sobre la geometría, consiguieron nuevos resultados sintéticamente. Los métodos geométricos, elegantes e intuitivamente claros, siempre cautivaron algunas mentes. Especialmente, Maclaurin prefirió la geometría sintética al análisis. La geometría pura, entonces, retuvo algo de vida aún si no se encontraba en el corazón de los desarrollos más vitales de los siglos XVII y XVIII. Al principio del siglo XIX varios grandes matemáticos decidieron que la geometría sintética había sido rechazada injusta y totalmente e hicieron un esfuerzo positivo por revivir y extender su enfoque.

Uno de los nuevos campeones de los métodos sintéticos, Jean-Victor Poncelet, concedió las limitaciones de la vieja geometría pura. Dice: «Mientras que la geometría analítica ofrece por su método general característico y uniforme medios de proceder a la solución

de las cuestiones que se nos presentan... mientras que llega a resultados cuya generalidad no tiene frontera, la otra [geometría sintética] procede por casualidad; su camino depende completamente de la sagacidad de aquellos que la emplean y sus resultados casi siempre están limitados a la figura particular que considera». Pero Poncelet no creía que los métodos sintéticos estuvieran necesariamente tan limitados, y propuso crear nuevos métodos que rivalizarían con el poder de la geometría analítica.

Michel Chasles (1793-1880) fue otro gran defensor de los métodos geométricos. En su *Aperçu historique sur l'origine et le développement des méthodes en géométrie (Resumen histórico sobre los orígenes y desarrollo de los métodos en geometría*, 1837), un estudio histórico en el cual Chasles admitió que ignoraba a los escritores alemanes porque desconocía su idioma, afirma que los matemáticos de su tiempo, y anteriores, habían declarado la geometría un lenguaje muerto que en el futuro no tendría uso ni influencia. No solamente Chasles niega esto, sino que cita a Lagrange, que era enteramente un analista, afirmando en su sesenta aniversario,[1] cuando había encontrado un problema muy difícil de mecánica celeste: «A pesar de que el análisis pueda tener ventajas sobre los viejos métodos geométricos, los cuales comúnmente, pero indebidamente, llamamos sintéticos, hay sin embargo problemas en el que el último aparece más ventajosamente, en parte debido a su claridad intrínseca y en parte debido a la elegancia y facilidad de sus soluciones. Hay aún algunos problemas para los cuales el análisis algebraico en alguna medida no es suficiente y que, según parece, sólo los métodos sintéticos pueden resolver.» Lagrange cita como ejemplo el muy difícil problema de la atracción que un elipsoide de revolución ejerce sobre un punto (unidad de masa) sobre su superficie o en el interior. Este problema había sido resuelto de modo puramente sintético por Maclaurin.

Chasles también proporciona un extracto de una carta que recibió de Lambert Adolphe Quetelet (1796-1874), el astrónomo y estadístico belga. Quetelet dice: «No es propio que la mayoría de nuestros matemáticos jóvenes valoren la geometría pura tan ligeramente.» Los jóvenes se quejaban de la carencia de generalidad del método, continúa Quetelet, pero es éste un fallo de la geometría o

[1] *Nouv. Mém. de l'Acad. de Berlin*, 1773, 121-148, pub. 1775 = *Œuvres*, 3, 617-658.

de aquellos que cultivan la geometría, se pregunta. Para compensar esta carencia de generalidad, Chasles proporciona dos reglas a los futuros geómetras. Deberán generalizar teoremas particulares para obtener los resultados más generales, que deberán ser al mismo tiempo simples y naturales. Segundo, no habrán de estar satisfechos con la demostración del resultado si no es parte de un método general o doctrina de la que depende naturalmente. Saber cuándo uno tiene realmente la base verdadera para un teorema, dice; siempre existe una verdad principal que uno reconocerá porque otros teoremas resultarán de la simple transformación, o como consecuencia inmediata. Las grandes verdades, que son el fundamento del conocimiento, siempre tienen las características de la simplicidad y la intuición.

Otros matemáticos atacaron los métodos analíticos con lenguaje más fuerte. Carnot deseaba «liberar a la geometría de los jeroglíficos del análisis». Más tarde, en el mismo siglo, Eduard Study (1862-1922) se refirió al proceso «mecánico» de la geometría con coordenadas como el «estruendo del molino coordenado».

Las objeciones a los métodos analíticos en geometría estuvieron basadas en algo más que una preferencia o gusto personal. Había, ante todo, la pregunta genuina de si la geometría analítica era realmente geometría, ya que el álgebra era la esencia del método y los resultados, y la significación geométrica de ambas estaba escondida. Más aún, señaló Chasles, el análisis, a través de sus procesos formales, niega todos los pequeños pasos que continuamente da la geometría. Los rápidos y tal vez penetrantes pasos del análisis no revelan el sentido de lo que se consigue. La conexión entre el punto inicial y el resultado final no es claro. Chasles pregunta: «¿Es entonces suficiente en un estudio filosófico y básico de una ciencia saber que algo es verdadero si uno no sabe por qué es así y qué lugar debería ocupar en la serie de verdades a lá que pertenece? El método geométrico, por otro lado, permite pruebas y conclusiones simples y evidentemente intuitivas.

Había otro argumento, el cual, mencionado primero por Descartes, aún gustaba en el siglo XIX. La geometría era considerada como la verdad sobre el espacio y el mundo exterior. El álgebra y el análisis no eran verdades significantes en sí mismas, ni siquiera acerca de números y funciones. Estos consistían únicamente en métodos para llegar a las verdades, y en eso se presentaban como artificiales. Esta visión del álgebra y del análisis estaba desapareciendo gradualmente. Sin embargo, la crítica era aún rigurosa al principio del si-

glo XIX, ya que los métodos del análisis eran incompletos y todavía sin base lógica. Los geómetras cuestionaban correctamente la validez de las pruebas analíticas y las aceptaban únicamente como sugiriendo resultados. El análisis sólo podía responder mordazmente que las demostraciones geométricas eran torpes y sin elegancia.

El resultado de la controversia es que los geómetras puros reafirmaban su papel en las matemáticas. Como si se tuvieran que vengar de Descartes porque su creación de la geometría analítica los había obligado a abandonar la geometría pura, los geómetras de principios del siglo XIX se propusieron como objetivo vencer a Descartes en su juego de la geometría. La rivalidad entre geómetras y analistas se hizo tan amarga que Steiner —un geómetra puro— amenazó con no publicar en el *Journal für Mathematik* de Crelle si éste continuaba publicando los artículos analíticos de Plücker.

El estímulo para revivir la geometría sintética vino principalmente de un solo autor, Gaspard Monge. Ya hemos discutido sus valiosas contribuciones a la geometría analítica y diferencial y sus inspiradoras clases en la Escuela Politécnica durante los años de 1795 a 1809. El mismo Monge no intentaba más que hacer regresar la geometría al rebaño de las matemáticas como un enfoque sugestivo y una interpretación de los resultados analíticos, y buscó únicamente el subrayar ambos modos de pensamiento. Sin embargo, su propio trabajo en geometría y su entusiasmo inspiró a sus alumnos Charles Dupin, François Joseph Servois, Charles Julien Brianchon, Jean Baptiste Biot (1744-1862), Lazare Nicholas Marguerite Carnot y Jean-Victor Poncelet, la urgencia por revitalizar la geometría pura.

La contribución de Monge a la geometría pura fue su *Traité de géométrie descriptive (Tratado de geometría descriptiva,* 1799). Esta materia muestra cómo proyectar ortogonalmente un objeto tridimensional en dos planos (uno horizontal y otro vertical) de tal forma que a partir de esta representación sea factible deducir propiedades matemáticas del objeto. El esquema es útil en arquitectura, diseño de fortificaciones, perspectiva, carpintería y talla de piedras, y fue el primero en tratar la proyección de una figura tridimensional en dos bidimensionales. Las ideas y métodos de la geometría descriptiva no demostraron ser una fuente de desarrollos posteriores en geometría o, en ese aspecto, en cualquier otra parte de las matemáticas.

2. Geometría euclídea sintética

A pesar de que los geómetras que reaccionaron a la inspiración de Monge se dedicaron a la geometría proyectiva, debemos detenernos para señalar algunos nuevos resultados en la geometría euclídea sintética. Estos resultados, tal vez pequeños en significación, muestran sin embargo nuevos temas y la casi infinita riqueza de esta vieja materia. De hecho, cientos de nuevos teoremas fueron obtenidos, de los cuales sólo podemos dar unos cuantos ejemplos.

Asociados con cada triángulo ABC están nueve puntos particulares, los puntos medios de los lados, los pies de las tres alturas y los puntos medios de los segmentos que unen los vértices con los puntos de intersección de las alturas. Los nueve puntos yacen sobre un círculo, llamado el círculo de los nueve puntos. Gergonne y Poncelet publicaron este teorema por primera vez.[2] Se le acredita frecuentemente a Karl Wilhelm Feuerbach (1800-1834), un maestro de escuela, quien publicó su demostración en *Eigenschaften einiger merkwürdigen Punkte des geradlinigen Dreiecks (Propiedades de algunos puntos distinguidos del triángulo rectilíneo,* 1822). En este libro, Feuerbach añadió otro hecho al círculo de los nueve puntos. un excírculo es uno que es tangente a uno de los lados y a las extensiones de los otros dos lados. (El centro de un excírculo yace sobre los bisectores de los dos ángulos exteriores y el ángulo interior más alejado.) El teorema de Feuerbach establece que el círculo de nueve puntos es tangente al círculo inscrito y a los tres excírculos.

En un pequeño libro publicado en 1816, *Über einige Eigenschaften des ebenen geradlinigen Dreiecks (Sobre algunas propiedades de triángulos rectilíneos planos),* Crelle mostró cómo determinar un punto P dentro de un triángulo tal que las líneas uniendo P a los vértices del triángulo y los lados del triángulo formen ángulos iguales. Esto es, $\sphericalangle 1 = \sphericalangle 2 = \sphericalangle 3$ en la figura 35.1. También existe un punto P' diferente de P tal que $\sphericalangle P'AC = \sphericalangle P'CB = \sphericalangle P'BA$.

Las secciones cónicas, como sabemos, fueron tratadas definitivamente por Apolonio como secciones de un cono y más tarde introducidas como lugares geométricos planos en el siglo XVII. En 1822, Germinal Dandelin (1794-1847) demostró[3] un teorema muy interesante acerca de las secciones cónicas en relación con un cono. Su

---

[2] *Ann. de Math.*, 11, 1820/21, 205-220.

[3] *Nouv. Mém. de l'Acad. Roy. des Sci.*, *Bruxelles*, 2, 1822, 169-202.

teorema establece que si dos esferas están inscritas en un cono circular de tal manera que son tangentes a un plano dado cortando al cono en una sección cónica, los puntos de contacto de las esferas con el plano son los focos de la sección cónica, y las intersecciones del plano con los planos de los círculos a lo largo de los que las esferas tocan el cono son las directrices de la cónica.

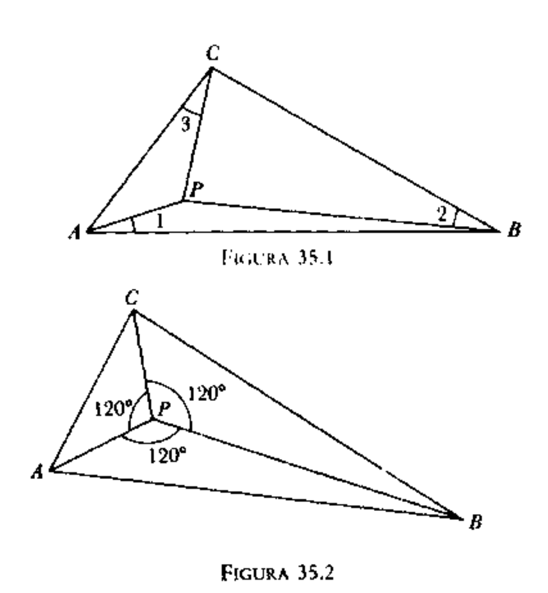


FIGURA 35.1

FIGURA 35.2

Otro tema interesante estudiado en el siglo XIX fue la solución de problemas de máximos y mínimos por métodos puramente geométricos, esto es, sin apoyarse en el cálculo de variaciones. De los varios teoremas que Jacob Steiner demostró usando únicamente métodos sintéticos, el resultado más famoso es el teorema isoperimétrico: de todas las figuras planas con un perímetro dado el círculo encierra el área máxima. Steiner proporcionó varias demostraciones.[4] Desafortunadamente, Steiner supuso que existe una curva que tiene área máxima. Dirichlet intentó varias veces persuadirlo de que sus

[4] *Jour. für Math.*, 18, 1838, 281-296; y 24, 1842, 83-162, 189-250; los artículos del 1842 se encuentran en su *Ges. Werke*, 2, 177-308.

demostraciones eran incompletas en ese punto pero Steiner insistía que esto era autoevidente. Una vez, sin embargo, escribió (en el primero de los ensayos de 1842):[5] «y la demostración se hace rápidamente si se supone que existe una figura máxima».

La demostración de la existencia de una curva maximizante frustró a los matemáticos por muchos años hasta que Weierstrass, en sus clases dadas durante los 1870 en Berlín, recurrió al cálculo de variaciones.[6] Más tarde, Constantin Caratheodory (1873-1950) y Study,[7] en un artículo conjunto, rigorizaron las demostraciones de Steiner sin emplear el cálculo. Sus demostraciones (hubo dos) fueron directas en lugar de indirectas, como en el método de Steiner. Hermann Amandus Schwarz, quien realizó una gran labor en ecuaciones en derivadas parciales y análisis y trabajó como profesor en diversas universidades, incluyendo Göttingen y Berlín, proporcionó una demostración rigurosa para el problema isoperimétrico en tres dimensiones.[8]

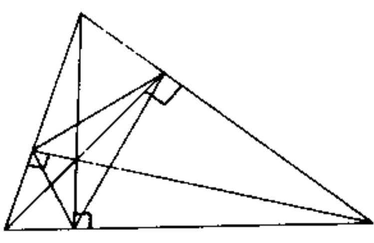


FIGURA 35.3

Steiner también demostró (en el primero de los ensayos de 1842) que de todos los triángulos con un perímetro dado, el equilátero tiene el área máxima. Otro de sus resultados [9] dice que si *A*, *B* ý *C* son tres puntos dados (fig. 35.2) y si cada uno de los ángulos del triángulo *ABC* es menor de 120°, entonces el punto *P* para el cual *PA* + *PB* + *PC* es un mínimo es tal que cada uno de los ángulos en

[5] *Ges. Werke*, 2, 197.

[6] *Werke*, 7, 257-264, 301-302.

[7] *Math. Ann.*, 68, 1909, 133-140 = Caratheodory, *Ges. Math. Schriften*, 2, 3-11.

[8] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1884, 1-13 = *Ges. Math. Abh.*, 2, 327-340.

[9] *Monatsber. Berliner Akad.*, 1837, 144 = *Ges. Werke*, 2, 93 y 729-731.

P es de 120°. Si, sin embargo, un ángulo del triángulo, digamos el ángulo *A*, es igual o más grande que 120°, entonces *P* coincide con *A*. Este resultado había sido demostrado mucho antes por Cavalieri (*Exercitationes Geometricae Sex*, 1647), pero indiscutiblemente le era desconocido a Steiner. Steiner también extendió este resultado a *n* puntos.

Schwarz resolvió el siguiente problema: dado un triángulo con ángulos agudos, considérense todos los triángulos tales que cada uno tiene sus vértices sobre los tres lados del triángulo original: el problema es encontrar el triángulo con el perímetro mínimo. Schwarz demostró sintéticamente [10] que los vértices de este triángulo de perímetro mínimo son los pies de las alturas del triángulo dado (fig. 35.3).[11]

En 1899, Frank Morley, profesor de matemáticas de la Universidad Johns Hopkins, descubrió un nuevo teorema de la geometría euclídea y demostraciones de él fueron publicadas por muchos autores.[12] El teorema dice que si son trazados los trisectores del ángulo en cada vértice de un triángulo, los trisectores adyacentes se cortan en los vértices de un triángulo equilátero (fig. 35.4). La novedad está en el uso de trisectores de ángulos. Hasta la mitad del siglo XIX

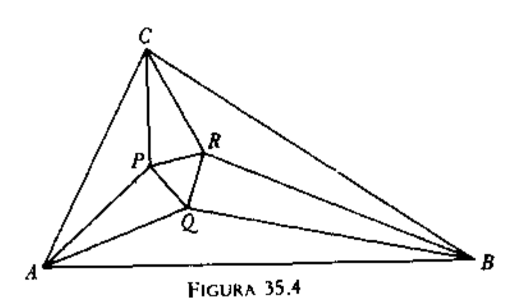


FIGURA 35.4

[10] Artículo inédito, *Ges. Math. Abh.*, 2, 344-345.

[11] La demostración de Schwarz se encuentra en Richard Courant y Herbert Robbins, *What is Mathematics?* Oxford University Press, 1941, 346-349. Una demostración usando el cálculo fue proporcionada por J. F. de Toschi di Fagnano (1715-1797) en el *Acta Eruditorum*, 1775, 297. Hubo demostraciones geométricas menos elegantes anteriores a la de Schwarz. Hay versión castellana, *¿Qué es la matemática?*, Madrid, Aguilar.

[12] Para una demostración y referencias a demostraciones publicadas véase: H. S. M. Coxeter, *Introduction to Geometry*, John Wiley and Sons, 1961, 23-25.

ningún matemático había considerado tales líneas ya que únicamente los elementos y figuras que son constructibles eran considerados con legitimidad en la geometría euclidiana. La constructividad garantizaba la existencia. Sin embargo, la concepción de lo que significaba la existencia cambió, como veremos más claramente cuando examinemos el trabajo sobre los fundamentos de la geometría euclidiana.

Algunos esfuerzos fueron dirigidos hacia la reducción del uso de regla y compás, en la misma línea iniciada por Mohr y Mascheroni (cap. 12, sec. 2). En su *Traité* de 1822, Poncelet demostró que todas las construcciones posibles con regla y compás (excepto la construcción de arcos circulares) eran posibles simplemente con regla, siempre que nos sea dado un círculo fijo con su centro. Steiner redemostró el mismo resultado más elegantemente en su pequeño libro *Die geometrischen Constructionen ausgeführt mittelst der geraden Linie und eines festen Kreises (Las construcciones geométricas ejecutadas por medio de regla y un círculo fijo).*[13] A pesar de que Steiner escribió el libro con fines pedagógicos, asegura en el prefacio del libro que establecerá la *conjetura* que un matemático francés había expresado.

La breve muestra anterior de teoremas euclídeos establecidos mediante métodos sintéticos no debe dejar al lector con la impresión de que no se usaron los métodos de la geometría analítica. De hecho, Gergonne proporcionó demostraciones analíticas de muchos teoremas geométricos que publicó en la revista fundada por él, los *Annales de Mathématiques*.

3. El resurgimiento de la geometría proyectiva sintética

El área principal hacia la que se volvieron Monge y sus alumnos fue la geometría proyectiva. Esta materia había tenido una explosión de actividad vigorosa, pero de breve vida, en el siglo XVII (cap. 14), pero fue menospreciada por el surgimiento de la geometría analítica, el cálculo y el análisis. Como ya hemos notado, el trabajo fundamental de Desargues de 1639 se perdió de vista hasta 1845 y el principal ensayo de Pascal sobre cónicas (1639) nunca fue recuperado. Sólo los libros de La Hire, que usaba algunos de los resultados de Desargues, estaban disponibles. Lo que los autores del siglo XIX

[13] Publicada 1833 = *Werke*, 1, 461-522.

aprendieron de los libros de La Hire, a menudo se lo atribuían a él incorrectamente. En general, estos geómetras ignoraban el trabajo de Desargues y Pascal y tuvieron que recrearlo. El resurgimiento de la geometría proyectiva fue iniciado por Lazare N. M. Carnot (1753-1823), alumno de Monge y padre del distinguido físico Sadi Carnot. Su principal obra fue la *Géométrie de position (Geometría de la posición,* 1803) y también contribuyó al *Essai sur la théorie des transversales (Ensayo sobre la teoría de las transversales,* 1806). Monge defendió el uso conjunto del análisis y la geometría pura, pero Carnot rehusó emplear métodos analíticos y empezó la cruzada de la geometría pura. Muchas de las ideas que discutiremos en breve están sugeridas casi en su totalidad en el trabajo de Carnot. Así, el principio que Monge llamó de las relaciones contingentes, conocido también como principio de correlatividad, y más comúnmente como principio de continuidad, se encuentra ahí. Carnot, a fin de evitar que las figuras sean separadas para los distintos tamaños de los ángulos y las direcciones de las líneas, no usó los números negativos, que consideraba contradictorios, y sí introdujo en cambio un complicado esquema llamado «correspondencia de signos».

Entre los cultivadores de principios del siglo XIX de la geometría proyectiva mencionaremos únicamente a François Joseph Servois y Charles Julien Brianchon (1785-1864), quienes aplicaron aspectos de su trabajo a problemas militares. Aunque contribuyeron a la reconstrucción, sistematización y extensión de viejos resultados, el único nuevo teorema importante es el famoso resultado de Brianchon,[14] que demostró mientras era estudiante en la Ecole. El teorema afirma que si tenemos seis tangentes a una cónica (fig. 35.5), formando un hexágono circunscrito, las tres rectas, cada una de las cuales une dos vértices opuestos, pasan por un punto. Brianchon derivó ese teorema utilizando la relación polo-polar.

El resurgimiento de la geometría proyectiva recibió su ímpetu principal de Poncelet (1788-1867). Poncelet fue alumno de Monge y también aprendió mucho de Carnot. Mientras servía como oficial en la campaña de Napoleón contra Rusia, es capturado y permanece el año 1813-1814 en una prisión rusa en Saratoff. Ahí Poncelet reconstruyó —sin la ayuda de libro alguno— lo que había aprendido de Monge y Carnot, creando nuevos resultados. Después amplió y revisó su trabajo, publicándolo bajo el título de *Traité des propriétés*

[14] Jour. de l'Ecole Poly., 6, 1806, 297-311.

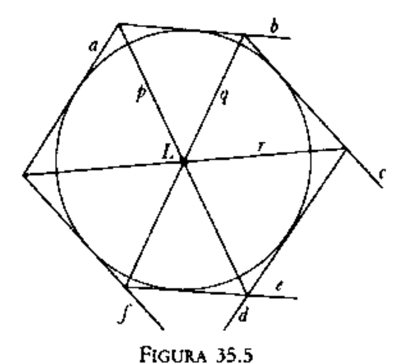


FIGURA 35.5

projectives des figures (Tratado de las propiedades proyectivas de las figuras, 1822). Este trabajo constituye su principal contribución a la geometría proyectiva y a la erección de una nueva disciplina. En su vida posterior se vio obligado a dedicar gran parte de su tiempo al servicio gubernamental, a pesar de que tuvo algunos nombramientos de profesor por períodos limitados.

Poncelet se convirtió en el más ardiente defensor de la geometría sintética y hasta atacó el análisis. Aunque había sido amigo del gran analista Joseph Díez Gergonne (1771-1859) y publicó ensayos en los *Annales de Mathématiques,* de Gergonne, sus ataques también fueron en breve dirigidos a Gergonne. Poncelet estaba convencido de la autonomía e importancia de la geometría pura. A pesar que admitía el poder del análisis, creía que era posible otorgar el mismo poder a la geometría sintética. En un ensayo de 1818, publicado en los *Annales* de Gergonne,[15] afirmó que el poder de los métodos analíticos no yacía en su uso del álgebra sino en su generalidad, y su ventaja resultaba del hecho que las propiedades métricas descubiertas para una figura típica permanecían aplicables, con la posible excepción del cambio de signo, a todas las figuras relacionadas que surgían de la típica o básica. Esta generalidad era asegurada en la geometría sintética por el principio de continuidad (que examinaremos en breve).

[15] *Ann. de Math.*, 8, 1817/18, 141-55. Este ensayo fue reimpreso en *Applications d'analyse et de geometrie* de Poncelet (1862-1864), 2, 466-476.

Poncelet fue el primer matemático en apreciar completamente que la geometría proyectiva era una nueva rama de las matemáticas, con métodos y metas propios. Mientras que los geómetras del siglo XVII habían tratado con problemas específicos, Poncelet entreveía el problema general de buscar todas las propiedades de las figuras geométricas que eran comunes a todas las secciones de cualquier proyección de una figura, esto es, permanecían sin alteración mediante la proyección y la sección. Este es el tema que él y sus sucesores estudiaron. Ya que las distancias y los ángulos son alterados por proyección y sección, Poncelet seleccionó y desarrolló la teoría de involución y de conjuntos armónicos de puntos, pero no el concepto de razón doble. Monge había utilizado la proyección paralela en su trabajo; como Desargues, Pascal, Newton y Lambert, Poncelet se valió de la proyección central, esto es, la proyección desde un punto, y elevó este concepto a un método de aproximación a los problemas geométricos. Poncelet también consideraba la transformación proyectiva a partir de una figura del espacio en otra, por supuesto, en una forma puramente geométrica. Aquí pareció perder interés en las propiedades proyectivas y estaba más interesado por el uso del método en bajo-relieve y diseño de andamios.

Su trabajo se centra en tres ideas. La primera es la de las figuras homólogas; dos figuras son homólogas si se deriva una de la otra mediante una proyección y una sección, que se denomina una perspectividad, o mediante una secuencia de proyecciones y secciones, esto es, una proyectividad. Al trabajar con figuras homólogas, su plan era encontrar para una figura dada una figura homóloga más simple, y, estudiándola, encontrar propiedades que son invariantes bajo proyección y sección, obteniendo propiedades de la figura más complicada. Desargues y Pascal emplearon la esencia de este método; y Poncelet, en su *Traité*, elogió la originalidad de Desargues en este y otros aspectos.

El segundo tema principal de Poncelet es el principio de continuidad. En su *Traité* lo establece de la siguiente manera: «Si una figura es derivada de otra mediante un cambio continuo y la última es tan general como la anterior, entonces cualquier propiedad de la primera figura puede ser establecida inmediatamente para la segunda.» La determinación de cuándo ambas figuras son generales no es explicada. El principio de Poncelet también afirma que si una figura degenera, como lo hace un hexágono en un pentágono cuando se hace que un lado se aproxime a cero, cualquier propiedad de la

figura original será transmitida con un argumento apropiadamente redactado para la figura degenerada.

El principio no era realmente nuevo con Poncelet. En su sentido filosófico amplio se remite a Leibniz, quien estableció en 1687 que cuando las diferencias entre dos casos se pueden hacer más pequeñas que cualquier dato dado, las diferencias se pueden hacer más pequeñas que cualquier cantidad dada en el resultado. Desde los tiempos de Leibniz el principio fue reconocido y usado constantemente. Monge empezó a valerse del principio de continuidad para establecer teoremas. Quería demostrar un teorema general, pero utilizó una posición especial de la figura para demostrarlo y entonces mantuvo que el teorema era cierto en general, aun cuando algunos elementos de las figuras se hicieran *imaginarios*. Así, para demostrar un teorema acerca de una recta y una superficie lo demostraba cuando la recta corta la superficie y mantiene entonces que el resultado es válido aun cuando la línea ya no corta a la superficie y los puntos de intersección son imaginarios. Ni Monge ni Carnot, quienes también aplicaron este principio, proporcionaron ninguna justificación.

Poncelet, quien acuñó el término «principio de continuidad», propuso el principio como una verdad absoluta y lo aplicó con libertad en su *Traité*. Para «demostrar» su solidez toma el teorema sobre la igualdad de los productos de los segmentos de cuerdas que se cortan en un círculo y nota que cuando el punto de intersección se mueve fuera del círculo se obtiene la igualdad de los productos de las secantes y sus segmentos externos. Más aún, cuando una secante se convierte en tangente, la tangente y su segmento externo se hacen iguales y su producto continúa siendo igual al producto de las otras secantes y su segmento externo. Todo esto era suficientemente razonable, pero Poncelet aplicó el principio para demostrar muchos teoremas y, como Monge, extendió el principio para hacer aserciones acerca de figuras imaginarias. (Mencionaremos algunos ejemplos más tarde.)

Los otros miembros de la Academia de Ciencias de París criticaron el principio de continuidad y lo consideraban como si tuviera únicamente un interés heurístico. Cauchy, en particular, criticó el principio, pero desafortunadamente la crítica fue dirigida hacia aplicaciones que Poncelet había realizado donde el principio sí funcionaba. Los críticos también mencionaron que la confianza que Poncelet y otros tenían en el principio provenía realmente del hecho de que podía ser justificado sobre una base algebraica. De hecho, las

notas que Poncelet hizo en prisión muestran que utilizó el análisis para probar la solidez de su principio. Estas notas, incidentalmente, fueron escritas por Poncelet y publicadas por él en dos volúmenes titulados *Applications d'analyse et de géométrie (Aplicaciones del análisis y la geometría*, 1862-1864), que es realmente una revisión de su *Traité* de 1822, y en el último trabajo sí usa métodos analíticos. Poncelet admitió que una demostración podía estar basada sobre el álgebra, pero insistió en que el principio no dependía de tal prueba. Sin embargo, es bastante seguro que Poncelet se apoyaba en el método algebraico para ver qué sucedía y entonces asegurar los resultados geométricos usando el principio como justificación.

Chasles, en su *Aperçu*, defendió a Poncelet. La posición de Chasles era que el álgebra es una demostración *a posteriori* del principio. Sin embargo, él se protegía señalando que había que ser cuidadoso en no llevar de una figura a otra cualquier propiedad que dependa *esencialmente* de que los elementos sean reales o imaginarios. De esta forma, una sección de un cono puede ser una hipérbola, y ésta tiene asíntotas. Cuando la sección es una elipse, las asíntotas se hacen imaginarias. De aquí que no era posible demostrar un resultado acerca de las asíntotas por sí solas, porque éstas dependían de la naturaleza particular de la sección. Como no se debían llevar los resultados para una parábola al caso de la hipérbola, porque el plano que corta no tiene una posición *general* en el caso de la parábola. Más adelante discute el caso de dos círculos intersecándose que tienen una cuerda en común. Cuando los círculos ya no se intersecan, la cuerda común es imaginaria. El hecho de que la cuerda común real pase por dos puntos *reales* es, dice, una propiedad incidental o contingente. Una cuerda se ha de definir en algún modo que no dependa del hecho de que pasa por puntos reales cuando los círculos se intersecan, sino que sea una propiedad permanente de los dos círculos en cualquier posición. Así, uno puede definirlo como el eje radical (real), que significa que es una recta tal que a partir de cualquier punto sobre ella, las tangentes a los dos círculos son iguales, o bien se puede definir por medio de la propiedad de que cualquier círculo trazado con cualquier punto de la recta como centro corta los dos círculos ortogonalmente.

Chasles también insistía en que el principio de continuidad era el adecuado para tratar los elementos imaginarios en geometría. Primero explica lo que se quiere decir por imaginario en geometría. Los elementos imaginarios pertenecen a una condición o estado de una

figura en el cual ciertas partes son no existentes, siempre que estas partes sean reales en otro estado de la figura. Porque, añade, uno no puede tener ninguna idea de las cantidades imaginarias, excepto pensando en los estados relacionados en que las cantidades son reales. Estos últimos estados son los que uno llama «accidentales» y los que proporcionan la clave para lo imaginario en geometría. Para demostrar resultados acerca de los elementos imaginarios sólo es necesario tomar la condición general de la figura en la que los elementos son reales, y entonces, de acuerdo con el principio de relaciones accidentales o el principio de continuidad, se concluye que los resultados se mantienen cuando los elementos son imaginarios. «Así se observa que el uso y la consideración de los imaginarios está completamente justificada.» El principio de continuidad fue aceptado durante el siglo XIX como intuitivamente claro y, por tanto, considerado como un axioma. Los geómetras lo usaron libremente y nunca juzgaron que requiriera una demostración.

A pesar de que Poncelet utilizó el principio de continuidad para aseverar resultados acerca de los puntos y líneas imaginarios, nunca proporcionó una definición general de tales elementos. Para introducir algunos puntos imaginarios proporcionó una definición geométrica que es complicada y nada clara. Entenderemos estos elementos imaginarios con más facilidad cuando sean discutidos desde un punto de vista algebraico. A pesar de la falta de claridad en el acercamiento de Poncelet, a él debe acreditársele la introducción de la noción de los puntos circulares en el infinito, esto es, dos puntos imaginarios situados sobre la recta en el infinito, y comunes a dos círculos cualesquiera.[16] También presentó los círculos esféricos imaginarios que dos esferas cualesquiera tienen en común. Más adelante demostró que dos cónicas reales que no se intersecan tienen *dos* cuerdas comunes imaginarias y dos cónicas se intersecan en cuatro puntos, reales o imaginarios.

La tercera idea conductora del trabajo de Poncelet es la noción de polo y polar con respecto a una cónica. El concepto data de Apolonio y fue usado por Desargues (cap. 14, sec. 3) y por otros en los trabajos del siglo XVII sobre geometría proyectiva: Euler, Legendre, Monge, Servois y Brianchon ya lo habían utilizado. Pero Poncelet proporcionó una formulación general de la transformación de polo a polar, y conversamente, y la usó en su *Traité* de 1822 y en

---

[16] *Traité*, 1, 48.

su «Mémoire sur la théorie générale des polaires réciproques» («Memoria sobre la teoría general de polares recíprocas»), presentada a la Academia de París en 1824 [17] como un método para establecer muchos teoremas.

Uno de los objetivos de Poncelet al estudiar la polar recíproca con respecto a una cónica era establecer el principio de dualidad. Los que trabajaban en geometría proyectiva habían observado que los teoremas relativos a figuras situadas en un plano, cuando eran parafraseados reemplazando la palabra «punto» por «recta» y «recta» por «punto», no solamente tenían sentido, sino que eran ciertos. La razón para la validez de los teoremas resultantes de tal paráfrasis no era clara, y de hecho Brianchon cuestionó el principio. Poncelet pensó que la relación entre polo y polar era la razón.

Sin embargo, esta relación requería la mediación de una cónica. Gergonne [18] insistió en que el principio era general y aplicable a todos los enunciados y teoremas, excepto aquellos que involucraban propiedades métricas. No se necesitaba del polo y la polar como un objeto de apoyo intermedio. Introdujo el término «dualidad» para denotar la relación entre el teorema original y el nuevo. También observó que en situaciones tridimensionales el punto y el plano constituyen elementos duales y la recta es dual es sí misma.

Para ilustrar el entendimiento por Gergonne del principio de dualidad, examinemos la dualización del teorema del triángulo de Desargues. Debemos notar primero lo que es el dual de un triángulo. Un triángulo consiste de tres puntos no todos sobre la misma recta, y las rectas que los unen. La figura dual consiste de tres rectas no todas ellas pasando por el mismo punto y los puntos uniéndolas (los puntos de intersección). La figura dual es de nuevo un triángulo y, por lo mismo, el triángulo es llamado autodual. Gergonne inventó el esquema de escribir teoremas duales en columnas dobles, con el dual a un lado de la proposición original.

Ahora consideremos el teorema de Desargues, donde esta vez los dos triángulos y el punto 0 yacen en un plano, y veamos qué resulta cuando intercambiamos punto por recta. En el ensayo de Gergonne, de 1825-1826, ya escribió este teorema y su dual como sigue:

---

[17] *Jour. für Math.*, 4, 1829, 1-71.

[18] *Ann. de Math.*, 16, 1825-1826, 209-231.

Teorema de Desargues	*Teorema Dual de Desargues*
Si tenemos dos triángulos tales que las rectas uniendo vértices correspondientes pasan por el punto O, entonces los lados correspondientes se intersecan en tres puntos sobre una recta.	Si tenemos dos triángulos tales que los puntos que son las uniones de lados correspondientes yacen sobre una recta O, entonces los vértices correspondientes están unidos por tres rectas pasando por el mismo punto.

Aquí el teorema dual es el recíproco del teorema original.

La formulación de Gergonne del principio general de dualidad fue algo vaga y tenía deficiencias. A pesar de que estaba convencido de que era un principio universal, no podía justificarlo y Poncelet señaló correctamente las deficiencias. También disputó con Gergonne sobre la prioridad del descubrimiento del principio (que realmente pertenecía a Poncelet) y aún acusó a Gergonne de plagio. Sin embargo, Poncelet se apoyaba realmente sobre el polo y la polar y no aceptaría el que Gergonne hubiera dado un paso adelante al reconocer las amplias aplicaciones del principio. Discusiones posteriores entre Poncelet, Gergonne, Möbius, Chasles y Plücker clarificaron el principio para todos y Möbius, en su *Der barycentrsiche Calcul (El Cálculo Baricéntrico),* más tarde, proporcionó un buen enunciado de la relación del principio de dualidad con el polo y la polar: la noción de dualidad es independiente de las cónicas y formas cuadráticas pero coincide con el polo y la polar cuando se usan las últimas. La justificación lógica del principio general de dualidad no fue proporcionada en este momento.

Jacob Steiner (1796-1863) llevó más lejos el desarrollo sintético de la geometría proyectiva. Es el primero de la escuela de geómetras alemanes que adoptó ideas francesas, notablemente de Poncelet y favoreció los métodos sintéticos hasta el punto de llegar a odiar el análisis. Hijo de un granjero suizo, trabajó también en la granja hasta la edad de 19 años. Aunque en gran parte era autodidacta, finalmente se convirtió en profesor en Berlín. En sus años mozos trabajó como maestro en la escuela de Pestalozzi y se impresionó ante la importancia que revestía incrementar la intuición geométrica. El principio de Pestalozzi era hacer que el estudiante creara las matemáticas con la guía del maestro y el método socrático. Steiner se fue a los extremos. Enseñó geometría pero no usaba figuras, y al

preparar a los candidatos al doctorado obscurecía la sala. En su propio trabajo posterior, Steiner se dedicó a tomar teoremas y demostraciones publicadas en inglés y en otras revistas y no daba indicaciones en sus propios textos respecto a que sus resultados ya habían sido publicados. Había hecho buen trabajo original durante su juventud y buscó mantener su reputación de productividad.

Su principal trabajo, *Systematische Entwicklung der Abhängigkeit geometrischen Gestalten von einander (Desarrollo sistemático de la dependencia de las formas geométricas una de otra,* 1832) y su objetivo principal fue usar los conceptos proyectivos para construir estructuras más complicadas a partir de las más simples como puntos, rectas, haces de rectas, planos y haces de planos. Sus resultados no eran especialmente nuevos, pero sí su método.

Para ilustrar su principio vamos a examinar su propio método proyectivo, ahora común, de definir las secciones cónicas. Se empieza (fig. 35.6) con dos haces de rectas (familias de rectas concurren-

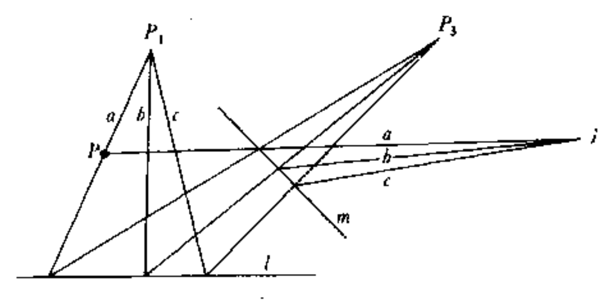


FIGURA 35.6

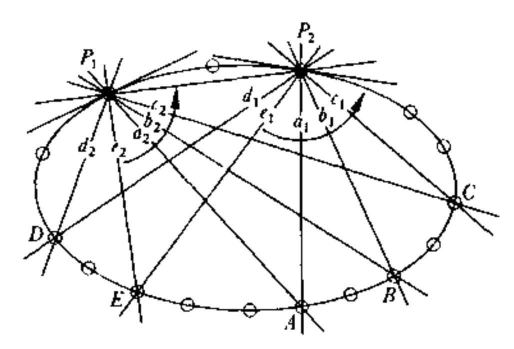


FIGURA 35.7

tes), digamos p_1 y p_3, que están relacionadas perspectivamente a través de un haz de puntos sobre una línea *l*, y los haces p_3 y p_2, que están relacionados perspectivamente por medio de un haz de puntos sobre otra línea *m*. Entonces se dice que los haces p_1 y p_2 están relacionados proyectivamente. Las líneas marcadas *a* en el haz con centro en P_1 y el haz con centro en P_2 son ejemplos de líneas correspondientes en la proyectividad entre los dos haces p_1 y p_2. Una cónica está definida ahora como el conjunto de los puntos de intersección de todos los pares de líneas correspondientes de los dos haces proyectivos. De esta manera, *P* es un punto sobre la cónica. Más aún, la cónica pasa por P_1 y P_2 (fig. 35.7). De esta manera, Steiner construyó las cónicas o curvas de segundo grado por medio de formas más simples, haces de rectas. Sin embargo, no identificó sus cónicas con las secciones de un cono.

También construyó las cuádricas regladas, el hiperboloide de una hoja y el paraboloide hiperbólico, de una manera semejante, haciendo de la correspondencia proyectiva la base de sus definiciones. De hecho, su método no es suficientemente general para toda la geometría proyectiva.

En sus demostraciones utilizó la razón doble como una herramienta fundamental, ignoró los elementos imaginarios y se refirió a ellos como los «fantasmas» o las «sombras de la geometría». Tampoco utilizó cantidades con signo, a pesar de que Möbius, cuyo trabajo examinaremos en breve, ya las había introducido.

Steiner utilizó el principio de dualidad desde el inicio de su trabajo. De tal forma, dualizó la definición de una cónica para obtener una nueva estructura llamada una curva de rectas. Si se empieza con dos haces de puntos relacionados proyectivamente (pero no perspectivamente), entonces la familia de líneas (fig. 35.8) uniendo los puntos correspondientes de los dos haces es llamada una cónica de rectas. Para distinguir tales familias de rectas, que también describen una curva, la curva usual como lugar geométrico de puntos es llamada curva de puntos. Las líneas tangentes a una curva de puntos son una curva de rectas y en el caso de una cónica constituyen la curva dual. Recíprocamente, cada cónica de rectas envuelve una cónica de puntos o es la colección de tangentes a una cónica de puntos.

Con la noción de Steiner de dual de una cónica de puntos, es posible dualizar muchos teoremas. Tomemos el teorema de Pascal y formemos el enunciado dual. Escribiremos el teorema a la izquierda y el nuevo enunciado a la derecha.

Teorema de Pascal

Si tomamos seis puntos *A*, *B*, *C*, *D*, *E* y *F* sobre una cónica de puntos, entonces las líneas que unen a *A* y *B* y *D* y *E* se juntan en el punto *P*; las líneas que unen *B* y *C* y *E* y *F* se juntan en el punto *Q*; las líneas que unen *C* y *D* y *F* y se juntan en el punto *R*. Los tres puntos *P*, *Q* y *R* yacen sobre una línea *l*.

Teorema dual de Pascal

Si tomamos seis líneas *a*, *b*, *c*, *d*, *e* y *f* sobre la cónica de rectas entonces los puntos que unen a *a* con *b* y *d* y *e* están unidos por la línea *p*; los puntos que unen *b* y *c* y *e* y *f* están unidos por la línea *q*; los puntos que unen *c* y *d* y *f* y *a* están unidos por la línea *r*. Las tres líneas *p*, *g* y *r* están sobre un punto *L*.

La figura 14.12 del capítulo 14 ilustra el teorema de Pascal. El teorema dual es el que Brianchon descubrió mediante la relación entre polo y polar (fig. 35.5 anterior). Steiner, como Gergonne, no hizo nada para establecer las bases lógicas del principio de dualidad. Sin embargo, desarrolló la geometría proyectiva sistemáticamente, clasificando las figuras y notando los argumentos duales a medida que avanzaba. También estudió en detalle las curvas y superficies de segundo grado.

Michel Chasles, quien dedicó su vida por completo a la geometría, siguió el trabajo de Poncelet y Steiner, aunque no conoció personalmente el trabajo de Steiner porque, como ya hemos notado, Chasles no sabía leer alemán. Chasles presentó sus propias ideas en su *Traité de géométrie supérieure* (*Tratado de geometría superior*, 1852) y el *Traité des sections coniques* (*Tratado de las secciones có-*

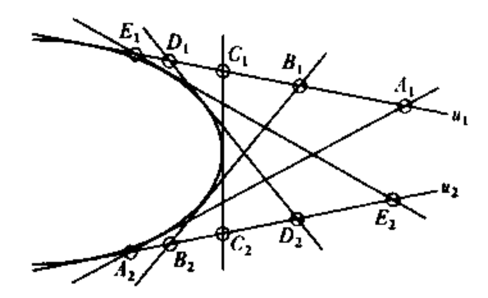


FIGURA 35.8

nicas, 1865). Ya que gran parte del trabajo de Chasles inintencionalmente duplicaba el de Steiner o fue superado por conceptos más generales, notaremos sólo unos cuantos resultados importantes que se le deben.

Chasles llegó a la idea de razón doble a partir de sus intentos por entender el trabajo perdido de Euclides *Los Porismas* (a pesar de que Steiner y Möbius ya lo habían reintroducido). Desargues también se había valido ya del concepto, pero Chasles conocía únicamente lo que La Hire había publicado sobre él. Chasles supo en algún momento que la idea también se encuentra en Pappus, ya que en la Nota IX de su *Aperçu* (p. 302) se refiere al uso de Pappus de la idea. Uno de los resultados de Chasles en este área [19] es que cuatro puntos fijos de una cónica y un quinto punto de la cónica determinan cuatro rectas con la misma razón doble.

En 1828, Chasles [20] dio el teorema: dados dos conjuntos de puntos colineales en una correspondencia uno-a-uno y tales que la razón doble de cuatro puntos cualesquiera sobre una recta es igual a la de los puntos correspondientes sobre la otra, entonces las rectas uniendo los puntos correspondientes son tangentes a una cónica que toca las dos rectas dadas. Este resultado es equivalente a la definición de Steiner de una cónica de rectas, ya que la condición de razón doble asegura aquí que los dos conjuntos de puntos colineales están relacionados proyectivamente y las rectas uniendo los puntos correspondientes son las rectas de la cónica de rectas de Steiner.

Chasles señaló que en virtud del principio de dualidad, las rectas pueden ser tan fundamentales como los puntos en el desarrollo de la geometría proyectiva plana y reconoce a Poncelet y Gergonne que son claros en este punto. Chasles introdujo también nueva terminología. A la razón doble la llamó proporción anharmónica. Introdujo el término homografía para describir una transformación de un plano en sí mismo u otro plano, que lleva puntos en puntos y rectas en rectas. Este término cubre figuras relacionadas proyectivamente u homólogas. Añadió la condición de que la transformación debe preservar la razón doble, pero este último hecho puede ser demostrado. A la transformación que lleva puntos en rectas y rectas en puntos la llamó correlación.

A pesar de que Chasles defendió la geometría pura, pensaba ana-

[19] *Correspondance mathématique et physique,* 5, 1829, 6-22.
[20] *Correspondance mathématique et physique,* 4, 1828, 363-371.

líticamente, aunque sus demostraciones y resultados los presentaba geométricamente. Este enfoque es llamado el «método mixto» y fue usado posteriormente por otros.

Hacia 1850 los conceptos y metas generales de la geometría proyectiva como distintos de los de la geometría euclidiana eran claros; sin embargo, la relación lógica entre las dos geometrías no estaba clara. El concepto de longitud había sido usado en geometría proyectiva de Desargues a Chasles. De hecho, el concepto de razón doble fue definido en términos de longitud. Sin embargo, la longitud no es un concepto proyectivo, porque no es invariante bajo las transformaciones proyectivas. Karl Georg Christian von Staudt (1798-1867), profesor en Erlangen, quien estaba interesado en los fundamentos, decidió liberar a la geometría proyectiva de su dependencia de la longitud y la congruencia. La esencia de su plan, presentada en su *Geometrie der Lage (Geometría de Posición*, 1847) consistió en introducir un análogo de la longitud sobre una base proyectiva. Su esquema es llamado «el álgebra de lanzamientos». Uno elige tres puntos arbitrarios sobre una línea y les asigna los símbolos 0, 1 y ∞. Entonces, por medio de una construcción geométrica (que de hecho proviene de Möbius) —un «lanzamiento»—, un símbolo es asociado a cualquier punto arbitrario P.

Para ver qué significa la construcción en geometría euclídea supongamos que empezamos con los puntos etiquetados 0 y 1 sobre la línea (fig. 35.9). Por el punto M sobre la línea paralela trácese $0M$ y entonces se traza $1N$ paralela a $0M$. Ahora se traza $1M$ y luego $N2$ paralela a $1M$. Es inmediato que $01 = 12$ ya que los lados opuestos de un paralelogramo son iguales. Entonces se traslada la longitud 01 a 12 mediante una construcción geométrica.

Ahora, para tratar el caso proyectivo empezamos con tres puntos: 0, 1, ∞ (fig. 35.10). El punto ∞ yace sobre l_∞, la recta del infinito, pero esta es una recta ordinaria en la geometría proyectiva. Ahora escójase un punto M y trácese una «paralela» a la recta 01 pasando por M. Esto significa que la recta que pasa por M debe cortar la recta 01 en ∞ y así trazamos $M\infty$. Ahora trazamos $0M$ y la prolongamos hasta l_∞. Después trazamos 1, una «paralela» a $0M$. Esto significa que la «paralela» que pasa por 1 debe cortar $0M$ sobre l_∞. Así obtenemos la recta $1\ P$ y esto determina N. Ahora se traza $1M$ y la prolongamos hasta que toca l_∞ en Q. La recta que pasa por N y es paralela a $1M$ es QN y donde corta a 01 obtenemos un punto que es etiquetado 2.

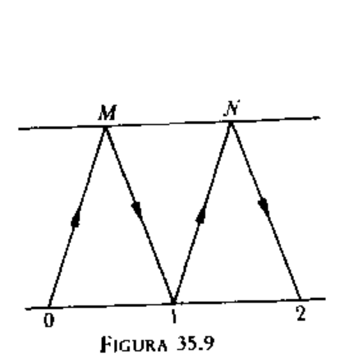


FIGURA 35.9

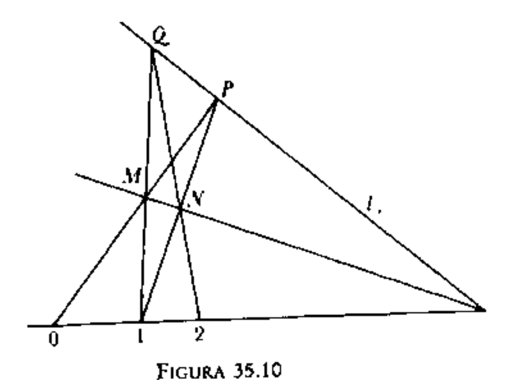


FIGURA 35.10

Por medio de este tipo de construcciones podemos asignar «coordenadas racionales» a los puntos sobre la recta 01∞. Para asignar números irracionales a puntos sobre la recta se debe introducir un axioma de continuidad (cap. 41). Esta noción no era bien entendida en aquel tiempo y, como consecuencia, al trabajo de Staudt le faltó rigor.

La manera como Von Staudt asignó las coordenadas a los puntos sobre la recta no hacía uso de la longitud. Sus coordenadas, a pesar de que eran los símbolos números usuales, sirvieron como símbolos de identificación sistemática para los puntos. Para añadir o sustraer tales «números», Von Staudt no podía usar las leyes de la aritmética. En su lugar, proporcionó construcciones geométricas que definían las operaciones con estos símbolos, por ejemplo, de tal forma que la suma de los números 2 y 3 es el número 5. Estas operaciones obedecían todas las leyes usuales de los números. Así, sus símbolos o coordenadas eran susceptibles a ser tratados como números ordinarios, a pesar de que habían sido construidos geométricamente.

Con estas etiquetas adheridas a los puntos, von Staudt podía definir la razón doble de cuatro puntos. Si las coordenadas de estos puntos son x_1, x_2, x_3 y x_4, entonces la razón doble está definida como

$$\frac{x_1 - x_3}{x_1 - x_4} \bigg/ \frac{x_2 - x_3}{x_2 - x_4}.$$

De este modo, von Staudt tenía las herramientas fundamentales para construir la geometría proyectiva sin depender de las nociones de longitud y congruencia.

Un conjunto armónico de cuatro puntos es uno para el cual la razón doble es −1. Sobre la base de los conjuntos armónicos von Staudt aportó la definición fundamental de que dos haces de puntos están relacionados proyectivamente cuando bajo una correspondencia uno-a-uno de sus elementos, a un conjunto armónico corresponde un conjunto armónico. Cuatro rectas concurrentes forman un conjunto armónico si los puntos en los que cortan a una transversal arbitraria constituyen un conjunto armónico de puntos. Así puede ser definida también la proyectividad de dos haces de líneas. Con estas nociones von Staudt definió una colineación del plano en sí mismo como una transformación uno-a-uno de punto a punto y línea a línea y demostró que lleva un conjunto de elementos armónicos sobre un conjunto de elementos armónicos.

La principal contribución de von Staudt en su *Geometrie der Lage* fue demostrar que la geometría proyectiva es claramente más fundamental que la euclídea. Sus conceptos son lógicamente anteriores. Este libro y su *Beiträge zur Geometrie der Lage (Contribuciones a la geometría de la posición*, 1856, 1857, 1860) revelaba la geometría proyectiva como una materia independiente de la distancia. Sin embargo, utilizó el axioma de las paralelas de la geometría euclídea, lo que desde un punto de vista lógico era una mancha, porque el paralelismo no es un invariante proyectivo. Esta mancha fue suprimida por Félix Klein.[21]

4. Geometría proyectiva algebraica

Mientras los geómetras sintéticos desarrollaban la geometría proyectiva, los geómetras algebraicos seguían tratando la misma materia con sus métodos. La primera de las nuevas ideas algebraicas fue la que ahora se denomina coordenadas homogéneas. Tal esquema fue creado por August Ferdinand Möbius (1790-1868) quien, como Gauss y Hamilton, vivía de la astronomía pero dedicó considerable tiempo a las matemáticas. Aunque Möbius no tomó parte en la controversia de los métodos sintéticos contra los algebraicos, sus contribuciones estuvieron del lado algebraico.

Su esquema para representar los puntos de un plano mediante

[21] *Math. Ann.*, 6, 1873, 112-145 = *Ges. Abh.*, 1, 311-343. También véase el cap. 38, sec. 3.

coordenadas, introducido en su trabajo principal, *Der barycentrische Calcul*,[22] consistió en empezar con un triángulo fijo y tomar como coordenadas de cualquier punto *P* en el plano la cantidad de masa que debe ser colocada en cada uno de los tres vértices de un triángulo, de tal forma que *P* sea el centro de gravedad de las tres masas. Cuando *P* está fuera del triángulo, entonces una o dos de las coordenadas pueden ser negativas. Si las tres masas son multiplicadas por la misma constante, el punto *P* sigue siendo el centro de gravedad. De aquí que sólo las proporciones de los tres están determinadas. El mismo esquema aplicado a puntos en el espacio requiere de cuatro coordenadas. Al escribir las ecuaciones de las curvas y superficies en este sistema coordenado, las ecuaciones se hacen homogéneas; esto es, todos los términos tienen el mismo grado. Veremos ejemplos del uso de coordenadas homogéneas en breve.

Möbius distinguió los tipos de transformaciones de un plano o espacio en otro. Si las figuras correspondientes son iguales, la transformación es una congruencia, y si son semejantes, la transformación es una semejanza. Siguiendo en generalidad viene la transformación que preserva el paralelismo, pero no la longitud o la forma y este tipo es llamado afín (noción introducida por Euler). La transformación más general que lleva rectas en rectas la llamó una colineación. Möbius demostró en su *Der barycentrische Calcul* que cada colineación es una transformación proyectiva; esto es, que resulta de una sucesión de perspectividades. Su demostración suponía que la transformación es uno-a-uno y continua, pero la condición de continuidad puede ser reemplazada por una más débil. También proporcionó una representación analítica de la transformación. Como señaló Mö-

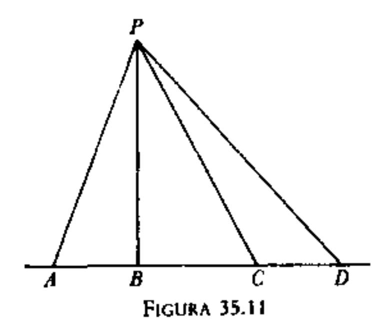


FIGURA 35.11

[22] Publicado en 1827 = *Werke*, 1, 1-388.

bius, se podían considerar propiedades invariantes de las figuras bajo cada uno de los tipos anteriores de transformaciones.

Möbius introdujo elementos con signo en geometría no solamente para los segmentos lineales sino también para las áreas y volúmenes. Consecuentemente, fue capaz de dar un tratamiento completo de la noción de razón doble con signo de cuatro puntos sobre una recta. Asimismo, demostró que la razón doble de cuatro rectas de un haz es posible expresarla en términos de los senos de los ángulos en el vértice P (fig. 35.11) mediante

$$\frac{\operatorname{sen} APB}{\operatorname{sen} APC} \Big/ \frac{\operatorname{sen} BPD}{\operatorname{sen} CPD}$$

y esta proporción es la misma que la razón doble de los cuatro puntos A, B, C y D cortados sobre cualquier transversal a las rectas del haz. De aquí que la razón doble quede inalterada por proyección y sección. Möbius tenía muchas otras ideas que desarrolló lentamente y no llevó muy lejos.

El hombre que dio al enfoque algebraico de la geometría proyectiva su eficacia y vitalidad es Julius Plücker (1801-1868). Después de ser profesor de matemáticas en varias instituciones desde 1836, se convirtió en profesor de matemáticas y física en Bonn, posición que mantuvo por el resto de su vida. Plücker era esencialmente un físico, y de hecho un físico experimental, en cuya actividad realizó muchos notables descubrimientos. A partir de 1863 se dedicó de nuevo a las matemáticas.

Plücker introdujo también las coordenadas homogéneas, pero de una manera diferente a Möbius. Su primera noción, las coordenadas trilineales,[23] fue presentada en el segundo volumen de su *Analytisch-geometrische Entwickelungen (Desarrollo de la geometría analítica*, 1828 y 1831). Empieza con un triángulo fijo y toma las coordenadas de cualquier punto P como las distancias perpendiculares con signo de P a los lados del triángulo; cada distancia puede ser multiplicada por la misma constante arbitraria. Más adelante, en el segundo volumen, introdujo un caso especial que consiste en considerar una recta del triángulo como recta del infinito. Esto es equivalente a reemplazar las coordenadas cartesianas rectangulares usuales x e y por $x = x_1/x_3$ e $y = x_2/x_3$ de tal forma que las ecuaciones de las

[23] *Jour. für Math.*, 5, 1830, 1-36 = *Wiss. Abh.*, 1, 124-158.

curvas se hacen homogéneas en x_1, x_2 y x_3. La última noción es la que fue usada más extensamente.

Mediante el uso de las coordenadas homogéneas y el teorema de Euler sobre funciones homogéneas, que establece que si $f(tx_1, tx_2, tx_3) = t^n f(x_1, x_2, x_3)$ entonces

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} = nf(x_1, x_2, x_3),$$

Plücker fue capaz de dar elegantes representaciones algebraicas de ideas geométricas. Así, si $f(x_1, x_2, x_3) = 0$ es la ecuación de una cónica con (x_1, x_2, x_3) como coordenadas de un punto sobre la cónica, entonces

$$\frac{\partial f}{\partial x_1} x'_1 + \frac{\partial f}{\partial x_2} x'_2 + \frac{\partial f}{\partial x_3} x'_3 = 0$$

puede ser interpretado, cuando x'_1, x'_2 y x'_3 son las coordenadas «libres», como la ecuación de la tangente en el punto (x_1, x_2, x_3) o, cuando x_1, x_2, x_3 son las coordenadas «libres», como la ecuación de la recta polar de un punto arbitrario (x'_1, x'_2, x'_3) con respecto a la cónica.

Usando coordenadas homogéneas, Plücker proporcionó la formulación algebraica de la recta infinitamente distante, los puntos circulares en el infinito y otras nociones. En el sistema de coordenadas homogéneas (x_1, x_2, x_3), la ecuación de la recta infinitamente distante es $x_3 = 0$. Esta recta no es excepcional en geometría proyectiva, pero en nuestra visualización de los elementos geométricos cada punto normal del plano euclídeo está en una posición finita dada por $x = x_1/x_3$ y $y = x_2/x_3$ y de esta manera estamos obligados a pensar en los puntos sobre $x_3 = 0$ como infinitamente distantes.

La ecuación de un círculo

$$(x - a)^2 + (y - b)^2 = r^2$$

se convierte con la introducción de coordenadas cartesianas homogéneas x_1, x_2 y x_3 a través de $x = x_1/x_3$ y $y = x_2/x_3$ en

$$(x_1 - ax_3)^2 + (x_2 - bx_3)^2 = r^1 x_3^2.$$

Ya que la ecuación de la línea infinitamente distante es $x_3 = 0$, la intersección de esta línea con el círculo está dada por

$$x_1^2 + x_2^2 = 0 \text{ y } x_3 = 0,$$

y esta es la ecuación de los puntos circulares en el ∞. Tales puntos circulares tienen las coordenadas $(1, i, 0)$ y $(1, -i, 0)$ o tripletas correspondientes a ellos. De la misma manera la ecuación de un círculo esférico *(Kugelkreis)* en el infinito es

$$x_1^2 + x_2^2 + x_3^2 = 0 \text{ y } x_1 = 0.$$

Si escribimos la ecuación de una línea recta en la forma homogénea (debemos usar x, y y z en lugar de x_1, x_2 y x_3)

$$Ax + By + Cz = 0$$

y ahora se requiere que la recta pase por los puntos (x_1, y_1, z_1) y $(1, i, 0)$ entonces la ecuación no homogénea resultante de la recta es

$$x - x_0 + i(y - y_0) = 0$$

donde $x_0 = x_1/z_1$ e $y_0 = = z_1/z_1$. De la misma manera, la ecuación de la recta que pasa por (x_1, y_1, z_1) y $(1, -i, 0)$ es

$$x - x_0 - i(y - y_0) = 0.$$

Cada una de estas rectas es perpendicular a sí misma, ya que la pendiente es igual a su recíproco negativo. Sophus Lie las llamó rectas locas; ahora son llamadas rectas isótropas.

Los esfuerzos de Plücker para tratar la dualidad algebraicamente lo llevaron a una idea muy bella, las coordenadas de rectas.[24] Si la ecuación de la recta en coordenadas homogéneas es

$$ux + vy + wz = 0,$$

y si x, y y z son cantidades fijas, u, v y w o cualesquiera tres números proporcionales a ellos son las coordenadas de una recta en un

[24] *Jour. für Math.*, 6, 1830, 107-146 = *Wiss. Abh.*, 1, 178-219.

plano. Entonces, de la misma manera que la ecuación $f(x_1, x_2, x_3) = 0$ representa una colección de puntos, así $f(u, v, w) = 0$ representa una colección de rectas o una curva reglada.

Con esta noción de coordenadas de rectas, Plücker fue capaz de proporcionar una formulación algebraica y una demostración del principio de dualidad. Dada cualquier ecuación $f(r,s,t) = 0$, si interpretamos r, s y t como las coordenadas homogéneas x_1, x_2 y x_3 de un punto, entonces tenemos la ecuación de una curva de puntos, mientras que si los interpretamos como u, v y w tenemos el dual de una curva de rectas. Cualquier propiedad demostrada mediante procesos algebraicos para la curva de puntos será, debido a que el álgebra es la misma bajo cualquier interpretación de las variables, origen de la propiedad dual de la curva de rectas.

En este segundo ensayo de 1830 y en el volumen 2 de su *Entwickelungen*, Plücker señaló también que la misma curva considerada como una colección de puntos puede asimismo ser considerada como la colección de rectas tangentes a la curva, ya que las tangentes determinan el contorno igual que lo hacen los puntos. La familia de las tangentes es una curva de rectas y tiene una ecuación en coordenadas de rectas. El grado de esta ecuación es llamado la clase de la curva, mientras que el grado de la ecuación en coordenadas de puntos es llamado el orden.

5. Curvas planas de orden superior y superficies

Los autores del siglo XVIII habían realizado algunos trabajos sobre curvas de grado superior al segundo (cap. 23, sec. 3) pero la materia estuvo latente desde 1750 a 1825. Plücker estudió curvas de grado tercero y cuarto y utilizó libremente conceptos proyectivos en su trabajo.

En su *System der analytischen Geometrie (Sistema de geometría analítica,* 1834) usó un principio que, aunque muy útil, difícilmente estaba bien fundamentado, para establecer formas canónicas de curvas. Para mostrar que la curva general de orden (grado) 4, digamos, podía ser puesta en una forma canónica, argumentó en particular que si el número de constantes fuera el mismo en las dos formas entonces se podía convertir una en la otra. Así, argumentó que una forma ternaria (tres variables) de cuarto orden podía siempre ser puesta en la forma

$$C_4 = pqrs + \mu\Omega^2,$$

donde p, q, r y s son formas lineales y Ω es una forma cuadrática, ya que ambos miembros contienen 14 constantes. En sus ecuaciones, μ era real, así como los coeficientes.

Plücker también estudió el número de puntos de intersección de curvas, un tema que también había sido considerado en el siglo XVIII. Utilizó un esquema para representar la familia de todas las curvas que pasan por las intersecciones de dos curvas C' y C'' de grado n-ésimo que había sido presentado por Lamé en un libro de 1818. Cualquier curva C_n pasando por estas intersecciones se podía expresar como

$$C_n = C'_n + \lambda C''_n = 0$$

donde λ es un parámetro.

Utilizando este esquema, Plücker proporcionó una clara explicación de la paradoja de Cramer (cap. 23, sec. 3). Una curva C_n general está determinada mediante $n(n + 3)/2$ puntos, ya que este es el número de coeficientes esenciales en su ecuación. Por otro lado, como dos C_n se intersecan en n^2 puntos, por los $n(n + 3)/2$ puntos en los cuales las dos C_n se intersecan, pasa un número infinito de otras C_n. La contradicción aparente fue explicada por Plücker.[25] Dos curvas cualesquiera de grado n se cortan, por supuesto, en n^2 puntos. Sin embargo, únicamente $(n/2)(n + 3) - 1$ puntos son independientes. Esto es, si tomamos dos curvas de grado n, que pasan por los $(n/2)(n + 3) - 1$ puntos, cualquier otra curva de grado n que pase, por estos puntos pasará por los restantes $(n - 1)(n - 2)/2$ de los n^2 puntos. Así, cuando $n = 4$, 13 puntos son independientes. Cualesquiera dos curvas que pasan por los 13 puntos determinan 16 puntos, pero cualquier otra curva que pase por los 13 pasará también por los tres restantes.

Plücker atacó después[26] la teoría de las intersecciones de una curva de grado m y una curva de grado n. Consideró la última como fija y las curvas intersecándola como variables. Usando la notación abreviada C_n para la expresión para la curva de n-ésimo grado y notación similar para las otras escribió

---

[25] *Annales de Math.*, 19, 1828, 97-106 = *Wiss. Abh.*, 1, 76-82.
[26] *Jour. für Math.*, 16, 1837, 47-54.

$$C_m = C'_m + A_{m-n}C_n = 0$$

para el caso donde $m > n$, de tal forma que A_{m-n} es un polinomio de grado $m - n$. A partir de esta ecuación obtuvo el método correcto para determinar las intersecciones de una C_n con todas las curvas de grado m. Ya que, de acuerdo con esta ecuación, $m - n + 1$ curvas linealmente independientes (el número de los coeficientes en A_{m-n}) pasan por las intersecciones de C'_m y C_n, la conclusión de Plücker fue que, dados $mn - (n - 1)(n - 2)/2$ puntos arbitrarios sobre C_n, los restantes $(n - 1)(n - 2)/2$ de los mn puntos de intersección con una C_m están determinados. El mismo resultado fue obtenido casi al mismo tiempo por Jacobi.[27]

En su *System* de 1834 y más explícitamente en su *Theorie der algebraischen Curven (Teoría de curvas algebraicas,* 1839), Plücker proporcionó las que ahora son llamadas fórmulas de Plücker, que relacionan el orden n y la clase k de una curva y las singularidades simples. Sea d el número de puntos dobles (puntos singulares en los cuales las dos tangentes son diferentes) y r el número de cúspides. A los puntos dobles les corresponde en la curva de rectas tangentes dobles (una tangente doble es de hecho tangente en dos puntos distintos) el número de las cuales es, digamos, t. A las cúspides les corresponde tangentes osculatrices (tangentes que cruzan la curva en puntos de inflexión) el número de las cuales es, digamos, w. Entonces Plücker pudo establecer las siguientes fórmulas duales

$$k = n(n - 1) - 2d - 3r \qquad n = k(k - 1) - 2t - 3w$$
$$w = 3n(n - 2) - 6d - 8r \qquad r = 3k(k - 2) - 6t - 8w.$$

El número de cualquier elemento incluye los casos reales e imaginarios.

En el caso cuando $n = 3$, $d = 0$ y $r = 0$, entonces w, el número de puntos de inflexión, es 9. Hasta los tiempos de Plücker, De Gua y Maclaurin habían demostrado que una recta que pasa por dos puntos de inflexión de la curva general de tercer grado pasa por un tercero, y el hecho de que una C_2 general tiene tres puntos reales de inflexión había sido supuesto desde los tiempos de Clairaut. En su *System* de 1834 Plücker demostró que una C_3 tiene uno o tres puntos reales de inflexión, y en el último caso estos están sobre una

[27] *Jour. für Math.*, 15, 1836, 285-308 = *Werke*, 3, 329-354.

recta. También llegó al resultado más general que toma en consideración elementos complejos. Una C_3 general tiene nueve puntos de inflexión de los cuales seis son imaginarios. Para derivar este resultado utilizó su principio de contar constantes para demostrar que

$$C_3 = fgh - l^3,$$

donde f, g, h y l son formas lineales, y obtuvo el resultado de De Gua y Maclaurin. Más adelante demostró (con argumentos incompletos) que los nueve puntos de inflexión de C_3 yacen sobre una recta de tal forma que hay doce de tales rectas. Ludwig Otto Hesse (1811-1874), quien trabajó como profesor en diversas universidades, completó la demostración de Plücker [28] y mostró que las doce rectas pueden ser agrupadas en cuatro triángulos.

Como un ejemplo adicional de descubrimiento de propiedades generales de las curvas, consideraremos el problema de los puntos de inflexión de una curva $f(x,y) = 0$ de n-ésimo grado. Plücker había expresado las condiciones comunes del cálculo para un punto de inflexión cuando $y = f(x)$, a saber, $d^2y/dx^2 = 0$, en la forma apropiada para $f(x,y) = 0$ y obtuvo una ecuación de grado $3n - 4$. Ya que la curva original y la nueva curva deben tener $n(3n - 4)$ intersecciones, parecía que la curva original tenía $n(3n - 4)$ puntos de inflexión. Como este número era demasiado grande, Plücker sugirió que la curva con una ecuación de grado $3n - 4$ tiene un contacto tangencial con cada una de las n ramas infinitas de la curva original $f = 0$, de tal forma que $2n$ de los puntos comunes no son puntos de inflexión, y así obtuvo el número correcto, $3n(n - 2)$. Hesse aclaró este punto [29] valiéndose de coordenadas homogéneas; esto es, reemplazó x por x_1/x_3 e y por x_2/x_3 y usando el teorema de Euler sobre funciones homogéneas demostró que la ecuación de Plücker para los puntos de inflexión puede ser escrita como

$$H = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} = 0,$$

[28] *Jour. für Math.*, 28, 1844, 97-107 = *Ges. Abh.*, 123-135.
[29] *Jour. für Math.*, 41, 1851, 272-284 = *Ges. Abh.*, 263-278.

donde los subíndices denotan derivadas parciales. Esta ecuación es de grado $3(n-2)$ y así toca a la curva $f(x_1, x_2, x_3) = 0$ de n-ésimo grado en el número correcto de puntos de inflexión. El determinante es llamado el hessiano de f, una noción introducida por Hess.[30]

Entre otros, Plücker abordó la investigación de curvas cuárticas. Fue el primero en descubrir (*Theorie der algebraischen Curven*, 1839) que tales curvas contienen 28 tangentes dobles de las cuales ocho a lo más son reales. Jacobi demostró entonces [31] que una curva de grado n-ésimo tiene en general $n(n-2)(n^2-9)/2$ tangentes dobles.

En el trabajo en geometría algebraica también se consideraron figuras en el espacio. Aunque la representación de líneas rectas en el espacio ya había sido introducida por Euler y Cauchy, Plücker en su *System der Geometrie des Raumes (Sistema de geometría del espacio*, 1846) introdujo una representación modificada

$$x = rz + \varrho, y = sz + \sigma$$

en la que los cuatro parámetros r, ϱ, s y σ fijan la recta. Ahora las rectas pueden ser usadas para construir todo el espacio, ya que, por ejemplo, los planos no son más que colecciones de rectas, y los puntos son intersecciones de rectas. Plücker dijo entonces que si las rectas eran consideradas como los elementos fundamentales del espacio, el espacio es tetradimensional porque son necesarios cuatro parámetros para cubrir todo el espacio con rectas. La noción de un espacio tetradimensional de puntos la rechazó como demasiado metafísica. Que la dimensión depende del elemento espacial es una idea nueva.

El estudio de figuras en el espacio incluyó superficies de tercer y cuatro grado. Una superficie reglada está generada por una recta moviéndose de acuerdo con alguna ley. El paraboloide hiperbólico (superficie con una silla) y el hiperboloide de una hoja son ejemplos de ellas, como lo es el helicoide. Si una superficie de segundo orden contiene una recta, contiene además una infinidad de rectas y es reglada. (Debe entonces ser un cono, un cilindro, el paraboloide hiperbólico, o el hiperboloide de una hoja.) Sin embargo, esto no es cierto para las superficies cúbicas.

Como ejemplo de propiedad notable de las superficies cúbicas

---

[30] *Jour. für Math.*, 28, 1844, 68-96 = *Ges. Abh.*, 89-122.
[31] *Jour. für Math.*, 40, 1850, 237-260 = *Werke*, 3, 517-542.

tenemos el descubrimiento de Cayley en 1849 [32] de la existencia de exactamente 27 rectas sobre cada superficie de tercer grado. No todas son necesariamente reales, pero existen superficies para las que todas son reales. Clebsch proporcionó un ejemplo de 1871.[33] Estas rectas tienen propiedades especiales. Por ejemplo, cada una es cortada por otras diez. Mucho más trabajo fue dedicado al estudio de estas rectas sobre superficies cúbicas.

Entre los muchos descubrimientos relativos a superficies de cuarto orden, uno de ellos, que se debe a Kummer, merece atención especial. Kummer había trabajado con familias de rectas que representan rayos de luz y, considerando las superficies focales asociadas [34], fue llevado a introducir una superficie de cuarto grado (y de cuarta clase) con 16 puntos dobles y 16 planos dobles como superficie focal de un sistema de rayos de segundo orden. Esta superficie, conocida como superficie de Kummer, contiene como caso especial la superficie de onda de Fresnel, que representa el frente de onda de la luz propagándose en un medio anisótropo.

La labor sobre geometría proyectiva sintética y algebraica de la primera mitad del siglo XIX abrió un período brillante para las investigaciones geométricas de toda clase. Los geómetras sintéticos dominaron el período. Atacaron cada resultado nuevo para descubrir algún principio general, usualmente no demostrable geométricamente, pero del que sin embargo derivaron un torrente de consecuencias ligadas unas a otras, y al principio general. Afortunadamente, también fueron introducidos los métodos algebraicos y, como veremos, dominaron finalmente el campo. Sin embargo, debemos interrumpir la historia de la geometría proyectiva para considerar algunas nuevas creaciones revolucionarias que afectaron el trabajo posterior en geometría y, que de hecho, alteraron radicalmente la imagen de la matemática.

Bibliografía

Berzolari, Luigi: «Allgemeine Theorie der höheren ebenen algebraischen Kurven». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III C 4, 313-455.

[32] *Cambridge and Dublin Math. Jour.*, 4, 1849, 118-132 = *Math. Papers*, 1, 445-456.

[33] *Math. Ann.*, 4, 1871, 284-345.

[34] *Monatsber. Berlinger Akad.*, 1864, 246-260, 495-499.

Boyer, Carl B.: *History of Analytic Geometry,* Scripta Mathematica, 1956, caps. 8-9.
Brill, A., y Noether, M.: «Die Entwicklung der Theorie der algebraischen Functionen in älterer und neurerer Zeit», *Jahres. der Deut. Math.-Verein.,* 3, 1892/3, 109-566, 287-312 en particular.
Cajori, Florian: *A history of mathematics,* 2.ª ed., Macmillan, 1919, 289-302, 309-314.
Coolidge, Julian L.: *A treatise on the circle and the sphere,* Oxford University Press, 1916.
— *A history of geometrical methods,* Dover (reimpresión), 1963, Libro I, cap. 5 y libro II, cap. 2.
— *A history of the conic sections and quadric surfaces.* Dover (reimpresión), 1968.
— «The rise and fall of proyective geometry». *American Mathematical Monthly,* 41, 1934, 217-228.
Fano, G.: «Gegensatz von synthetischer und analytischer Geometrie in seiner historichen Entwicklung im XIX, Jahrhuundert», *Encyk. der Math. Wiss.,* B. G. Teubner, 1907-1910, III AB4a, 221-288.
Klein, Felix: *Elementary mathematics from an advanced standpoint,* Macmillan, 1939; Dover (reimpresión), 1945, Geometría, parte 2. Hay versión castellana, *Matemática elemental desde un punto de vista superior* (2 volúmenes), Madrid, Biblioteca Matemática, 1927 y 1931.
Kotter, Ernst: «Die Entwickelung der synthetischen Geometrie von Monge bis auf Staudt», 1847, *Jahres. der Deut. Math-Verein.,* vol. 5, parte II, 1896 (pub. 1901), 1-486.
Möbius, August F.: *Der barycentrische Calcul* (1827), Georg Olms (reimpresión), 1968. También en vol. 1 de *Gesammelte Werke,* 1-388.
Plücker, Julius: *Gesammelte wissenschaftliche Abhandlungen,* 2 vols., B. G. Teubner, 1895-1896.
Schoenflies, A.: «Projektive Geometrie», *Encyk. der Math. Wiss.,* B. G. Teubner, 1907-1910, III AB5, 389-480.
Smith, David Eugene: *A source book in mathematics,* Dover (reimpresión), 1959, vol. 2, 315-323, 331-345, 670-676.
Steiner, Jacob: *Geometrical constructions with a ruler* (una traducción de su libro de 1833), Scripta Mathematica, 1950.
— *Gesammelte Werke,* 2 vols., G. Reimer, 1881-1882; Chelsea (reimpresión), 1971.
Zacharias, M.: «Elementargeometrie and elementare nicht-euklidische Geometrie in synthetischer Behandlund», *Encyk. der Math. Wiss.,* B. G. Teubner, 1907-1910, III, AB9, 859-1172.

Capítulo 36
LA GEOMETRIA NO EUCLIDEA

... porque parece ser cierto que muchas cosas tienen, por así decirlo, una época en la que son descubiertas en muchos lugares simultáneamente, como las violetas aparecen en primavera por todas partes.

WOLFGANG BOLYAI

El carácter de necesidad adscrito a las verdades de la matemática, y aún la peculiar certeza a ellas atribuida, son una ilusión.

JOHN STUART MILL

1. Introducción

Entre todas las complejas creaciones técnicas del siglo XIX la más profunda, la geometría no euclídea, fue técnicamente la más simple. Esta creación dio origen a nuevas e importantes ramas de las matemáticas, pero su implicación más significativa es que obligó a los matemáticos a revisar radicalmente su comprensión de la naturaleza de las matemáticas y su relación con el mundo físico. Dio origen a problemas en los fundamentos de las matemáticas con los que el siglo XX sigue luchando. Como veremos, la geometría no euclídea fue la culminación de una larga serie de esfuerzos en el área de la geometría euclídea. El fruto de este trabajo llegó en la primera parte del siglo XIX, durante las mismas décadas en las que la geometría proyectiva estaba siendo revivida y extendida. Sin embargo, los dos dominios no estuvieron relacionados uno con otro en esta época.

2. El status de la geometría euclídea hacia 1800

A pesar de que los griegos habían reconocido que el espacio abstracto o matemático es distinto de las percepciones sensoriales del espacio, y Newton subrayó este punto,[1] hasta aproximadamente 1800 todos los matemáticos estaban convencidos de que la geometría euclídea era la idealización correcta de las propiedades del espacio físico y de las figuras en ese espacio. De hecho, como ya hemos notado, hubo muchos intentos por construir la aritmética, el álgebra y el análisis, cuyos fundamentos lógicos eran obscuros, sobre la geometría euclídea y, por ende, garantizar también la verdad de esas ramas.

De hecho, muchos proclamaron la verdad absoluta de la geometría euclídea. Por ejemplo, Isaac Barrow, quien construyó sus matemáticas (incluyendo su cálculo) sobre la geometría, da una lista de ocho razones de la certeza de la geometría: la claridad de sus conceptos, las definiciones no ambiguas, la seguridad intuitiva y la verdad universal de sus axiomas, la posibilidad clara y lo fácil de imaginar de sus postulados, el pequeño número de sus axiomas, la visión clara del modo en que las magnitudes son generadas, el fácil orden de sus demostraciones y la elusión de cosas no conocidas.

Barrow formuló la pregunta: ¿cómo estamos seguros de que los principios geométricos se aplican a la naturaleza? Su respuesta fue que éstos son derivados de la razón innata. Los objetos sensibles son únicamente los agentes que los despiertan. Más aún, los principios de la geometría han sido confirmados mediante la experiencia constante y continuará siendo así porque el mundo, diseñado por Dios, es inmutable. La geometría es entonces la ciencia perfecta y segura.

Es relevante que los filósofos de la última parte del siglo XVII y el XVIII también formulasen la pregunta de cómo podemos estar seguros de que la mayor parte del cuerpo del conocimiento que la ciencia newtoniana ha producido era cierto. Casi todos, pero notablemente Hobbes, Locke y Leibniz, contestaron que las leyes matemáticas, como la geometría euclídea, eran inherentes al diseño del universo. Leibniz dejó cierto espacio para la duda cuando distinguió entre mundos posibles y reales. Pero la única excepción significativa fue David Hume, quien en su *Treatise of human nature (Tratado sobre la naturaleza humana,* 1739), negó la existencia de leyes o

[1] *Principia,* Libro I, Def. 8, Escolio.

secuencias necesarias de eventos en el universo, y afirmó que se observaba que ocurrían estas secuencias y los seres humanos concluyeron que siempre ocurrirán de la misma manera. La ciencia es puramente empírica. En particular, las leyes de la geometría euclídea no son necesariamente verdades físicas.

La influencia de Hume fue negada y claramente superada por la de Immanuel Kant. La respuesta de Kant a la pregunta de cómo podemos estar seguros de que la geometría euclídea se aplica al mundo físico, que proporcionó en su *Crítica de la razón pura* (1781), es peculiar: Kant sostenía que nuestra mente suministra ciertos modos de organización —los llamó intuiciones— del espacio y el tiempo y que la experiencia es absorbida y organizada por nuestras mentes de acuerdo con esos modos o intuiciones. Nuestras mentes están de tal forma construidas que nos obligan a ver el mundo externo sólo de una manera. Como consecuencia, ciertos principios acerca del espacio son anteriores a la experiencia; estos principios y sus consecuencias lógicas, que Kant llamó verdades sintéticas *a priori*, son las de la geometría euclídea. Conocemos la naturaleza del mundo exterior sólo en la medida en que nuestras mentes nos obligan a interpretarla. Sobre las bases antes descritas, Kant afirmó, y sus contemporáneos aceptaron, que el mundo físico debía ser euclídeo. En cualquier caso, ya sea que se recurra a la experiencia, se apoye sobre verdades innatas, o acepte el punto de vista de Kant, la conclución común era la unidad y necesidad de la geometría euclídea.

3. Las investigaciones sobre el axioma de la paralelas

A pesar de que la confianza en la geometría euclídea como la idealización correcta del espacio físico permaneció firme del 300 a. de C. hasta cerca de 1800, una preocupación ocupó a los matemáticos durante casi todo ese período. Los axiomas adoptados por Euclides se suponía que eran verdades autoevidentes acerca del espacio físico y de las figuras en ese espacio. Sin embargo, el axioma de las paralelas en la forma establecida por Euclides (cap. IV, sec. 3) se pensaba que era algo demasiado complicado. Nadie dudaba realmente de su verdad, y sin embargo carecía de la cualidad obligada de los otros axiomas. Aparentemente, al propio Euclides no le gustaba su propia versión del axioma de las paralelas, ya que no no lo utilizó

sino después de que hubo demostrado todos los teoremas que pudo sin él.

Un problema relacionado que no preocupó a tanta gente pero que, sin embargo, finalmente llegó a ser igualmente vital es si se puede suponer la existencia de líneas rectas infinitas en el espacio físico. Euclides fue cuidadoso en postular que uno únicamente puede prolongar una línea recta (finita) tan lejos como sea necesario, de tal forma que aún la recta extendida seguía siendo finita. También la redacción peculiar del axioma de las paralelas de Euclides, que dos rectas se cortarán del lado de la transversal donde la suma de los ángulos interiores es menor que dos ángulos rectos, era una manera de evitar la aserción inequívoca de que existen pares de rectas que nunca se llegan a cortar por lejos que sean extendidas. Sin embargo, Euclides sugirió la existencia de líneas rectas infinitas ya que, de ser finitas, no podían ser extendidas tanto como fuera necesario en un contexto dado y *demostró* la existencia de rectas paralelas.

La historia de la geometría no euclídea se inicia con los esfuerzos para eliminar las dudas acerca del axioma de las paralelas de Euclides. Desde los tiempos de los griegos hasta aproximadamente 1800, se llevaron a cabo dos intentos. Uno fue reemplazar el axioma de las paralelas por un argumento más autoevidente. El otro fue tratar de deducirlo de los otros nueve axiomas de Euclides; de ser esto posible, sería un teorema y así estaría fuera de toda duda. No daremos una reseña completamente detallada de este trabajo porque su historia se encuentra fácilmente disponible.[2]

El primer intento importante fue hecho por Ptolomeo en su opúsculo sobre el axioma de las paralelas. Intentó demostrar esta aseveración deduciéndola de los otros nueve axiomas de Euclides y de los teoremas 1 a 28, que no dependen del axioma de las paralelas. Pero Ptolomeo asumió inconscientemente que dos líneas rectas no encierran un espacio y que si *AB* y *CD* son paralelas (fig. 36.1) entonces todo lo que se tenga para los ángulos interiores sobre un lado de *FG* debe mantenerse para el otro.

Proclo, el comentador del siglo V, fue muy explícito en cuanto a esta objeción al axioma de las paralelas. Dice: «Este [postulado] debería ser incluso ser eliminado de los postulados; ya que es un teorema que supone demasiadas dificultades, que Ptolomeo, en cierto libro, se propuso resolver, y requiere para su demostración un gran

---

[2] Véase, por ejemplo, la referencia de Bonola en la bibliografía al final del capítulo.

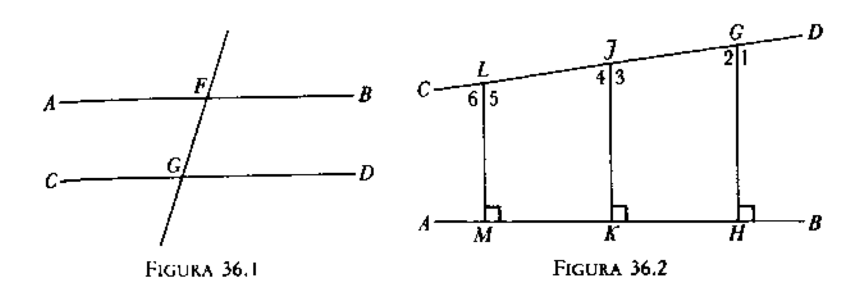


FIGURA 36.1 FIGURA 36.2

número de definiciones así como de teoremas; y el recíproco es, de hecho, demostrado por el propio Euclides como un teorema». Proclo señala que mientras que es necesario creer que dos líneas rectas tenderán una hacia la otra del lado de la transversal donde la suma de los ángulos interiores es menor que dos ángulos rectos, no es tan claro que estas dos rectas se intersecarán de hecho en un punto finito. Esta conclusión es únicamente probable. Ya que, continúa, hay ciertas curvas que se aproximan una a la otra más y más, pero que de hecho no se juntan. Así, una hipérbola se aproxima, pero no toca, a su asíntota. ¿No sería esto verdadero para las dos líneas en el postulado de Euclides? Más adelante dice que hasta cierta suma de los ángulos interiores de un lado de la transversal, las dos líneas deben realmente cortarse; sin embargo, para un valor escasamente más grande pero aún menor que dos ángulos rectos, las líneas podían ser asintóticas.

Proclo basó su propia demostración del postulado de las paralelas en un axioma que Aristóteles había usado para demostrar que el universo es finito. El axioma dice: «si desde un punto dos líneas rectas formando un ángulo son extendidas indefinidamente, las distancias sucesivas entre dichas líneas rectas [perpendiculares de una a la otra] excederán finalmente cualquier magnitud finita». La demostración de Proclo era correcta esencialmente, pero sustituyó un axioma cuestionable por otro.

Nasîr Eddîn (1201-1274), el editor persa de Euclides, proporcionó análogamente una «demostración» del postulado de las paralelas de Euclides suponiendo que dos líneas no paralelas se aproximan una a la otra en una dirección y divergen en la otra. Específicamente, si *AB* y *CD* (fig. 36.2) son dos rectas cortadas por *GH*, *JK*, *LM*, ..., si estas últimas son perpendiculares a *AB*, y si los ángulos 1, 3, 5,

... son obtusos mientras que 2, 4, 6, ... son agudos, entonces $GH > JK > LM$ ... Este hecho, afirmar Nasîr Eddîn, se ve claramente.

Wallis realizó cierto trabajo sobre el axioma de las paralelas en 1663, que publicó en 1693.[3] Primero, reprodujo el trabajo de Nasîr Eddîn sobre el axioma de las paralelas que había sido traducido para él por un profesor de árabe de Oxford. Incidentalmente, así fue como el trabajo de Nasîr-Eddîn sobre el axioma de las paralelas fue conocido en Europa. Entonces, Wallis criticó la demostración de Nasîr Eddîn y ofreció su propia demostración de la aserción de Euclides. Su demostración se apoya en la suposición explícita de que para cada triángulo existe uno semejante, cuyos lados están en una razón dada con los lados del original. Wallis pensaba que este axioma era más evidente que la subdivisión arbitrariamente pequeña y la extensión arbitrariamente grande. De hecho, dijo, el axioma de Euclides de que podemos construir un círculo con un centro y radio dados presupone que existe un radio arbitrariamente grande a nuestra disposición. De aquí que se pueda suponer el resultado análogo para figuras rectilíneas tales como un triángulo.

En 1769, Joseph Fenn sugirió el más sencillo de los axiomas substitutos, a saber, que dos líneas rectas intersecándose no pueden ser ambas paralelas a una tercera línea recta. Este axioma también aparece en los comentarios de Proclo sobre la proposición 31 del libro I de los *Elementos* de Euclides. El argumento de Fenn es equivalente por completo al axioma dado en 1795 por John Playfair (1748-1819): por un punto dado P que no está sobre una recta l, existe una sola recta en el plano de P y l que no corta a l. Este axioma se usa en los libros de texto modernos (que por simplicidad dicen comúnmente: «una y sólo una recta...»).

Legendre trabajó sobre el problema del postulado de las paralelas a lo largo de un período de veinte años. Sus resultados aparecieron en libros y artículos, incluyendo las muchas ediciones de sus *Eléments de géométrie (Elementos de Geometría)*.[4] En uno de los ataques al problema demostró el postulado de las paralelas suponiendo que existen triángulos semejantes de diferentes tamaños; de hecho esta demostración era analítica, pero supuso que la unidad de longitud no tenía importancia. Más adelante proporcionó una demostración basada en la suposición de que dados tres puntos cualesquie-

---

[3] *Opera*, 2, 669-678.

[4] 1.ª ed., 1794.

ra no colineales existe una circunferencia pasando por los tres. En otro enfoque usó todos los postulados, excepto el de las paralelas, y demostró que la suma de los ángulos de un triángulo no puede ser mayor que dos ángulos rectos. Entonces observó que bajo estas mismas suposiciones el área es proporcional al defecto, esto es, dos ángulos rectos menos la suma de los ángulos. Por tanto, trató de construir un triángulo de dos veces el tamaño de un triángulo dado de tal manera que el defecto del más grande fuera dos veces el del dado. Procediendo de esta manera, esperó obtener triángulos con defectos mayores y menores y obtener así sumas de ángulos aproximándose a cero. Este resultado, pensó, sería absurdo, y así la suma de los ángulos tendría que ser 180°. Este hecho, a su vez, implicaría el axioma euclídeo de las paralelas. Pero Legendre encontró que la construcción se reducía a demostrar que a través de cualquier punto dado dentro de un ángulo menor de 60°, uno siempre puede trazar una línea recta que toque a ambos lados del ángulo, y esto no lo podía demostrar sin hacer uso del postulado de las paralelas de Euclides. En cada una de las doce ediciones (12.ª ed., 1813) de la versión de Legendre de los *Elementos* de Euclides añadió apéndices que supuestamente proporcionaban demostraciones del postulado de las paralelas, pero cada uno era deficiente porque suponía algo implícito que no podía ser supuesto o bien aceptaba axiomas tan cuestionables como el de Euclides.

En el transcurso de sus investigaciones, Legendre,[5] usando los axiomas euclídeos con excepción del axioma de las paralelas, demostró los siguientes teoremas significativos: si la suma de los ángulos de un triángulo es dos ángulos rectos, entonces es así en cada triángulo. También, si la suma es menor que dos ángulos rectos en un triángulo, así es también en cada triángulo. Más adelante, proporciona la demostración de que si la suma de los ángulos de cualquier triángulo es dos ángulos rectos, se mantiene el postulado de las paralelas de Euclides. Este trabajo sobre la suma de los ángulos de un triángulo tampoco tuvo frutos, porque Legendre fue incapaz de demostrar (sin la ayuda del axioma de las paralelas o un axioma equivalente) que la suma de los ángulos de un triángulo no puede ser menos que dos ángulos rectos.

Los esfuerzos descritos hasta ahora fueron simplemente intentos por encontrar un substituto más autoevidente del axioma de las pa-

---

[5] *Mém. de l'Acad. des Sci., Paris,* 12, 1833, 367-410.

ralelas de Euclides y muchos de los axiomas propuestos parecían ser intuitivamente más autoevidentes. Consecuentemente, sus creadores estaban satisfechos, ya que habían cumplido sus objetivos. Sin embargo, un examen más detenido mostró que estos axiomas sustitutos no eran necesariamente más satisfactorios. Algunos de ellos hicieron aserciones acerca de lo que sucede indefinidamente lejos en el espacio. Así, requerir que hubiera un círculo pasando por tres puntos cualesquiera no en línea recta requiere círculos más y más grandes en cuanto los puntos se aproximan a estar alineados. Por otro lado, los axiomas sustitutos que no hacían intervenir el infinito directamente, por ejemplo, el axioma que establece que existen dos triángulos semejantes pero no iguales, se veían como suposiciones demasiado complejas y de ninguna manera preferibles al propio axioma de Euclides.

El segundo grupo de esfuerzos para resolver el problema del axioma de las paralelas buscaba deducir la aserción de Euclides de los otro nueve axiomas. La deducción podía ser directa o indirecta. Ptolomeo había intentado una prueba directa. El método indirecto supone una aserción contradictoria en lugar del enunciado de Euclides e intenta deducir una contradicción dentro del nuevo cuerpo de teoremas deducidos. Por ejemplo, ya que el axioma de las paralelas de Euclides es equivalente al axioma de que por un punto P que no está sobre una recta l hay una y solamente una paralela a l, existen dos alternativas a este axioma. Una es que no hay paralelas a l pasando por P y la otra es que existe más de una paralela a l pasando por P. Si al adoptar una de estas posibilidades en lugar del axioma de «una paralela» se demuestra que el nuevo conjunto conduce a una contradicción, entonces estas alternativas tendrían que ser rechazadas y la aserción de «una paralela» quedaría *demostrada*.

Gerolamo Saccheri (1667-1733), sacerdote jesuita y profesor de la universidad de Padua, realizó el esfuerzo más significativo en esta dirección. Estudió cuidadosamente los trabajos de Nasîr Eddîn y Wallis y adoptó su propio enfoque. Saccheri empezó con el cuadrilátero $ABCD$ (fig. 36.3) en el que A y B son ángulos rectos y $AC = BD$. Es fácil demostrar que $\sphericalangle C = \sphericalangle D$. Ahora, el axioma de Euclides es equivalente a la aserción de que los ángulos C y D son ángulos rectos. Entonces Saccheri consideró dos alternativas posibles:

1) la hipótesis del ángulo obtuso: $\sphericalangle C$ y $\sphericalangle D$ son obtusos;

2) la hipótesis del ángulo agudo: $\sphericalangle C$ y $\sphericalangle D$ son agudos.

Sobre la base de la primera hipótesis (y los otros nueve axiomas

de Euclides), Saccheri demostró que los ángulos *C* y *D* deben ser ángulos rectos. Así, de esta hipótesis dedujo una contradicción.

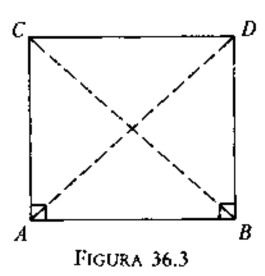


FIGURA 36.3

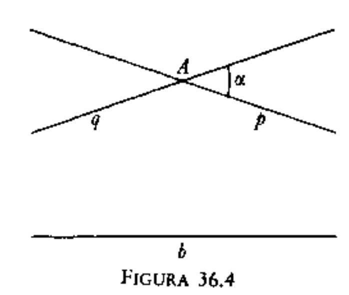


FIGURA 36.4

A continuación, Saccheri consideró la segunda hipótesis y demostró muchos teoremas interesantes. Continuó hasta que llegó al siguiente teorema: dado cualquier punto *A* y una recta *b* (fig. 36.4), con la hipótesis del ángulo agudo existe en el haz (familia) de rectas que pasan por *A* dos rectas *p* y *q* que dividen el haz en dos partes. La primera de éstas consiste de las líneas que intersecan a *b*, y la segunda consiste de las líneas (que forman un ángulo α) que tienen una perpendicular común con *b* en algún sitio a lo largo de *b*. Las rectas *p* y *q* ellas mismas son asintóticas a *b*. De este resultado y un argumento sumamente largo, Saccheri dedujo que *p* y *b* tendrían una perpendicular común en su punto común, el cual está en el infinito. A pesar de que no había obtenido contradicción alguna, Saccheri encontró esta conclusión, y otras, tan repugnantes que decidió que la hipótesis del ángulo agudo debía ser falsa.

Quedaba únicamente la hipótesis de que los ángulos *C* y *D* de la figura 36.3 eran ángulos rectos. Saccheri había demostrado previamente que cuando *C* y *D* son ángulos rectos, la suma de los ángulos de cualquier triángulo es igual a 180° y que este hecho implica el postulado de las paralelas de Euclides. Por tanto, se sintió justificado cuando concluyó su apoyo a Euclides y así publicó su *Euclides ab Omni Naevo Vindicatus (Euclides vindicado de todo reproche*, 1733). Sin embargo, puesto que Saccheri no obtuvo contradicción alguna sobre la base de la hipótesis del ángulo agudo, el problema del axioma de las paralelas permanecía abierto.

Los esfuerzos por encontrar un sustituto adecuado para el axioma euclídeo de las paralelas o demostrar que la aserción euclídea

debía ser un teorema fueron tan numerosas y tan fútiles que en 1759 d'Alembert llamó al problema del axioma de las paralelas «el escándalo de los elementos de la geometría».

4. Presagios de las geometrías no euclídeas

En su disertación de 1763, Georg S. Klügel (1739-1812), profesor de matemáticas de la universidad de Helmstadt, quien conocía el libro de Saccheri, observó sagazmente que la certeza con la que las gentes aceptaban la verdad del axioma euclidiano de las paralelas estaba apoyada en la experiencia. Esta observación introdujo por primera vez la idea de que la experiencia, por encima de la autoevidencia, sostenía los axiomas. Klügel expresó dudas acerca de que la aserción euclidiana pudiera ser demostrada. Se dio cuenta de que Saccheri no había llegado a contradicción alguna, sino simplemente a resultados que parecían estar en oposición con la experiencia.

El ensayo de Klügel le sugirió ideas sobre el axioma de las paralelas a Lambert. En su libro *Theorie der Parallellinien (Teoría de las líneas paralelas)* escrito en 1766, pero publicado en 1786,[6] Lambert, de alguna manera como Saccheri, consideró un cuadrilátero con tres ángulos rectos y consideró las posibilidades de que un cuarto ángulo fuese recto, obtuso, o agudo. Lambert descartó la hipótesis del ángulo obtuso porque lo llevó a una contradicción. Sin embargo, contrariamente a Saccheri, Lambert no concluyó que había obtenido una contradicción de la hipótesis del ángulo agudo.

Lambert dedujo de ambas hipótesis de los ángulos obtuso y agudo, respectivamente, dos consecuencias que es importante mencionar aunque la primera no le llevó a contradicción. Su resultado más sorprendente es que bajo cualquiera de las hipótesis el área de un polígono de n lados es proporcional a la diferencia entre la suma de los ángulos y $2n - 4$ ángulos rectos. (Saccheri poseía este resultado para un triángulo.) También notó que la hipótesis del ángulo obtuso originaba teoremas como los que se obtenían para figuras sobre la superficie de una esfera. Conjeturó que los teoremas que se seguían de la hipótesis del ángulo agudo se aplicarían a figuras sobre una esfera de radio imaginario. Esto lo condujo a escribir un ensayo[7]

---

[6] *Magazin für reine und angewandte Mathematik*, 1786, 137-164, 325-358.

[7] *Hist. de l'Acad. de Berlin*, 24, 1768, 327-354, pub. 1770 = *Opera Mathematica*, 2, 245-269.

sobre las funciones trigonométricas de ángulos imaginarios, esto es, iA, donde A es un ángulo real e $i = \sqrt{-1}$, lo que en efecto introducía las funciones hiperbólicas (cap. 19, sec. 2). Más tarde veremos un poco más claro el significado de las observaciones de Lambert.

Los puntos de vista de Lambert sobre la geometría eran bastante avanzados. Se dio cuenta que cualquier conjunto de hipótesis que no conducía a contradicciones ofrecía una geometría posible. Tal geometría sería una estructura lógica válida a pesar de que no tuviera relación con figuras reales. Lo último debe ser sugestivo para una geometría en particular pero no controlaba la variedad de geometrías desarrollables lógicamente. Lambert no llegó a la conclusión más radical, introducida un poco más tarde por Gauss.

Ferdinand Karl Schweikart (1780-1859), un profesor de jurisprudencia que dedicaba su tiempo libre a las matemáticas, dio otro paso hacia adelante. Trabajó sobre las geometrías no euclídeas durante el período en el que Gauss dedicó alguna atención a la materia, pero Schweikart llegó a sus conclusiones independientemente. Sin embargo, Saccheri y Lambert influenciaron sus trabajos. En un memorandum de 1816 que envió a Gauss en 1818 para su aprobación. Schweikart distinguió de hecho dos geometrías. Existía la geometría de Euclides y una geometría basada en la suposición que la suma de los ángulos de un triángulo no es igual a dos rectos. A esta última geometría la llamó geometría astral, porque podía cumplirse en el espacio de las estrellas, y sus teoremas eran los que Saccheri y Lambert habían establecido sobre la base de la hipótesis del ángulo agudo.

Franz Adolf Taurinus (1794-1874), un sobrino de Schweikart, siguió la sugerencia de su tío de estudiar la geometría astral. Aunque estableció en su *Geometriae Prima Elementa (Geometría de Elementos Primeros,* 1826) algunos nuevos resultados, sobre todo para la geometría analítica, concluyó que únicamente la geometría de Euclides podía ser verdadera para el espacio físico, pero que la geometría astral era *lógicamente consistente.* Taurinus demostró también que las fórmulas que eran verdaderas en esferas de radios imaginarios son precisamente aquellas que lo eran en su geometría astral.

Los trabajos de Lambert, Schweikart y Taurinus constituyen avances que merecen una recapitulación. Los tres, así como otros autores como Klügel y Abraham G. Kastner (1719-1800), profesor de Göttingen, estaban convencidos de que el axioma de las paralelas de Euclides no podía ser demostrado, esto es, que es independiente de los otros axiomas de Euclides. Más aún, Lambert, Schweikart y

Taurinus estaban convencidos de que es posible adoptar un axioma alternativo contradiciendo el de Euclides y construir una geometría lógicamente consistente. Lambert no hizo aserciones sobre la aplicabilidad de tal tipo de geometría; Taurinus pensaba que no era aplicable al mundo físico; pero Schweikart pensaba que se podía aplicar a la región de las estrellas. Estos tres autores también notaron que la geometría sobre una esfera real tiene las propiedades de la geometría basada en la hipótesis del ángulo obtuso (si se deja de lado la contradicción que resulta de la última) y que la geometría sobre la esfera de radio imaginario tiene las propiedades de una geometría basada en la hipótesis del ángulo agudo. Así, los tres reconocieron la existencia de la geometría no euclídea pero fallaron en un punto fundamental, a saber, en que la geometría euclídea no es la única geometría que describe las propiedades del espacio físico dentro de la precisión de la que la experiencia puede responder.

5. La creación de la geometría no euclídea

Ninguna de las ramas principales de la matemática, e incluso ningún resultado específico es obra de un solo hombre. A lo más, algún paso decisivo o demostración puede ser acreditada a un individuo. Este desarrollo acumulativo de la matemática se aplica especialmente a la geometría no euclídea. Si se interpreta la creación de la geometría no euclídea como el reconocimiento de que pueden existir geometrías alternativas a la de Euclides, entonces Klügel y Lambert merecen la cita. Si geometría no euclídea significa el desarrollo técnico de las consecuencias de un sistema de axiomas que contiene una alternativa al axioma de las paralelas de Euclides, entonces la mayor parte del mérito debe ser atribuido a Saccheri, y aún él se benefició del trabajo de muchos que intentaron encontrar un axioma sustitutivo más aceptable que el de Euclides. Sin embargo, el hecho más significativo acerca de la geometría no euclídea es que puede ser usada para describir las propiedades del mundo físico tan precisamente como lo hace la geometría euclídea. La última no es necesariamente la geometría del espacio físico; su verdad física no puede ser garantizada sobre una base *a priori*. Gauss logró esta comprensión por primera vez, comprensión que no requería de ningún desarrollo matemático técnico, puesto que éste ya había sido dado.

Carl Friedrich Gauss (1777-1855) fue hijo de un albañil de la

ciudad alemana de Brunswick y parecía estar destinado al trabajo manual. Pero el director de la escuela en la que recibió su educación elemental estaba impresionado por la inteligencia de Gauss y lo recomendó al duque Karl Wilhelm. El duque envió a Gauss a una escuela superior y, más adelante, en 1795, a la universidad de Göttingen. Ahí empezó Gauss a trabajar duro en sus ideas. A los 18 años inventó el método de mínimos cuadrados y a los 19 demostró que es posible construir un polígono regular de 17 lados. Estos éxitos lo convencieron de que debería cambiar de la filología a las matemáticas. En 1798 se mudó a la universidad de Helmstadt y ahí llamó la atención de Johann Friedrich Pfaff, quien se convirtió en su maestro y amigo. Después de terminar su doctorado, Gauss regresó a Brunswick donde escribió algunos de sus más famosos trabajos. Con este trabajo ganó, en 1807, el puesto de profesor de astronomía y director del observatorio de Göttingen. Residió en Göttingen por el resto de su vida, con excepción de una visita a Berlín para asistir a una reunión científica. No le gustaba enseñar y así lo afirmaba. Si embargo, disfrutaba de la vida social, se casó dos veces y creó una familia.

El primer trabajo importante de Gauss fue su tesis doctoral, en la que demostró el teorema fundamental del álgebra. En 1801 publicó su clásico *Disquisitiones Arithmeticae*. Su trabajo matemático en geometría diferencial, las «Disiquisitiones generales circa superficies curvas» («Investigaciones generales sobre superficies curvas», 1827), que, incidentalmente, fue el resultado de su interés en agrimensura, geodesia y cartografía, es una joya matemática (cap. 37, sec. 2). Realizó muchas otras contribuciones al álgebra, funciones complejas y teoría del potencial. Mantuvo inédito su trabajo innovativo en dos campos principales: las funciones elípticas y la geometría no euclídea.

Sus intereses en física fueron igualmente muy amplios y les dedicó gran parte de sus energías. Cuando Giuseppe Piazzi (1746-1826) descubrió el planeta Ceres en 1801, Gauss se dedicó a determinar su trayectoria. Este fue el inicio de su trabajo en astronomía, actividad que lo absorbió mucho y a la que dedicó cerca de veinte años. Una de sus grandes contribuciones en este área es su *Theoria Motus Corporum Colestium (Teoría del movimiento de los cuerpos celestes*, 1809). También se hizo merecedor de alta consideración por sus investigaciones físicas sobre magnetismo teórico y experimental. Maxwell comenta en su *Electricity and Magnetism (Electricidad y*

Magnetismo) que los estudios de magnetismo de Gauss reconstruyeron la ciencia por completo, en los instrumentos usados, los métodos de observación y el cálculo de los resultados. Los ensayos de Gauss sobre magnetismo terrestre son modelos de investigación física y proporcionaron el mejor método para medir el campo magnético de la tierra. Su trabajo en astronomía y magnetismo abrió un nuevo y brillante período de alianza entre las matemáticas y la física.

A pesar de que Gauss y Wilhelm Weber (1804-1891) no inventaron la idea de la telegrafía, en 1833 mejoraron técnicas anteriores con un aparato práctico que hacía rotar una aguja a derecha o izquierda dependiendo de la dirección de la corriente enviada a través del alambre. Gauss también trabajó en óptica, que había sido olvidada desde los tiempos de Euler, y sus investigaciones de 1838-1841 proporcionaron una base completamente nueva para el estudio de los problemas ópticos.

La universalidad de las actividades de Gauss es de lo más sorprendente, ya que sus contemporáneos habían empezado a encerrarse en investigaciones especializadas. A pesar del hecho de ser reconocido como el más grande de los matemáticos, al menos desde Newton, Gauss no fue tanto un gran innovador sino más bien una figura de transición entre los siglos XVIII y XIX. A pesar de llegar a algunos nuevos puntos de vista que influenciaron a otros matemáticos, estaba orientado más hacia el pasado que hacia el futuro. Felix Klein describe la posición de Gauss con estas palabras: nosotros podemos tener un cuadro del desarrollo de las matemáticas si nos imagináramos una cadena de altas montañas representando a los autores del siglo XVIII coronados por una cima imponente —Gauss— seguida de una larga y rica región llena de nuevos elementos de vida. Los contemporáneos de Gauss reconocieron su genio y, para su muerte, en 1855, era ampliamente venerado y llamado «el príncipe de los matemáticos».

Gauss publicó relativamente poco de su trabajo, ya que pulía cualquier cosa que hacía, en parte para lograr elegancia y en parte para lograr en sus demostraciones el máximo de brevedad sin sacrificar el rigor, al menos el rigor de su tiempo. En el caso de la geometría no euclídea no publicó ningún trabajo definitivo. En una carta a Bessel de 27 de enero de 1829, dijo que probablemente nunca publicaría sus hallazgos en esta materia porque temía al ridículo o, como dijo, temía al clamor de los beocios, una referencia figurativa a una tribu griega estúpida. Gauss pudo haber sido demasiado cui-

dadoso, pero se recordará que, a pesar de que algunos matemáticos habían llegado gradualmente al clímax de la geometría no euclídea, el mundo intelectual, en general, seguía dominado por las enseñanzas de Kant. Lo que conocemos del trabajo de Gauss en geometría no euclídea está tomado de las cartas a sus amigos, dos breves reseñas en el *Göttingische Gelehrte Anzeigen* de 1816 y 1822 y algunas notas de 1831 encontradas entre sus papeles después de su muerte.[8]

Gauss era completamente consciente de los vanos esfuerzos por establecer el postulado de las paralelas de Euclides, ya que esto era conocimiento común en Göttingen; y Kastner, el maestro de Gauss, estaba completamente familiarizado con la historia de estos esfuerzos. Gauss le comentó a su amigo Schumacher que, ya desde 1792 (por aquel entonces Gauss contaba quince años de edad) pensaba que podía existir una geometría lógica en la que el postulado de las paralelas de Euclides no se cumpliera. En 1794 Gauss había encontrado que, en su concepto de geometría no euclídea, el área de un cuadrángulo debía ser proporcional a la diferencia entre 360° y la suma de los ángulos. Sin embargo, en este fecha posterior y aún en 1799, Gauss insistía en deducir el postulado de las paralelas de Euclides a partir de otras suposiciones más plausibles y aún creía que la geometría euclídea era la geometría del espacio físico, y ello a pesar de que podía concebir otras geometrías lógicas no euclídeas. Sin embargo, el 17 de diciembre de 1799, Gauss escribió a su amigo el matemático húngaro Wolfgang Farkas Bolyai (1775-1856).

En cuanto a mí, he logrado cierto progreso en mi trabajo. Sin embargo, el camino que he escogido no lleva del todo a la meta que buscamos [la deducción del axioma de las paralelas], y el cual me aseguras que has logrado. Me parece mejor dudar de la verdad de la propia geometría. Es cierto que he logrado lo que la gente sostendría que constituye una demostración; pero a mis ojos no prueba nada. Por ejemplo, si pudiéramos demostrar que un triángulo rectilíneo cuya área pudiera ser más grande que cualquier área es posible, entonces estaría listo para demostrar la totalidad de la geometría [euclídea] de manera absolutamente rigurosa.

La mayoría de la gente ciertamente lo dejaría como un axioma, ¡pero yo no! Sería, por supuesto, posible que el área permanezca siempre menor que un cierto límite, por lejos que los tres puntos angulares del triángulo fueran tomados.

---

[8] *Werke*, 8, 157-268, contiene todos los anteriores y la carta discutida más abajo.

Este pasaje muestra que, en 1799, Gauss estaba hasta cierto punto convencido de que el axioma de las paralelas no podía ser deducido de los restantes axiomas euclídeos y comenzó a tomar más en serio el desarrollo de una nueva y posiblemente aplicable geometría.

A partir de 1813, Gauss desarrolló su nueva geometría, que primero llamó geometría antieuclídea, más adelante geometría astral y finalmente geometría no euclídea. Se convenció de que era lógicamente consistente y estaba bastante seguro de que podía ser aplicable. En reseñas de 1816 y 1822 y en una carta a Bessel de 1829, Gauss reafirmó que el postulado de las paralelas no podía ser demostrado sobre la base de los otros axiomas de Euclides. Su carta a Olbers escrita en 1817 [9] es un momento culminante. En ésta, Gauss afirma: «Me estoy convenciendo más y más de que la necesidad física de nuestra geometría euclídea no puede ser demostrada, al menos no por la razón humana ni para la razón humana. Tal vez en otra vida seamos capaces de obtener una visión de la naturaleza del espacio, que es ahora inalcanzable. Hasta entonces no debemos situar la geometría en la misma clase que la aritmética, que es puramente *a priori,* sino con la mecánica.»

Para poner a prueba la aplicabilidad de la geometría euclídea y su geometría no euclídea, Gauss midió de hecho la suma de los ángulos de un triángulo formado por los picos de tres montañas: Brocken, Hohehagen y Inselsberg. Los lados de los triángulos fueron 69, 85 y 197 km. Encontró [10] que la suma excedía 180° en 14″85. El experimento no demostró nada, ya que el error experimental era mucho más grande que el exceso y así la suma correcta podía haber sido 180° o aún menos. Como se dio cuenta Gauss, el triángulo era pequeño y como en la geometría no euclídea el defecto es proporcional al área, únicamente un triángulo grande podía posiblemente revelar alguna diferencia significativa de la suma de ángulos de 180°.

No discutiremos los teoremas particulares de la geometría no euclídea que se deben a Gauss. No escribió una presentación deductiva completa y los teoremas que demostró son muy parecidos a los que encontraremos en los trabajos de Lobatchevsky y Bolyai. Estos dos son considerados generalmente autores de la creación de la geometría no euclídea. Lo que se les tiene que atribuir será discutido posteriormente, pero ellos publicaron presentaciones organizadas de

---

[9] *Werke*, 8, 177.

[10] *Werke*, 4, 258.

una geometría no euclídea sobre una base sintética deductiva con el entendimiento completo de que esta nueva geometría era lógicamente tan legítima como la de Euclides.

Nikolai Ivanovich Lobatchevsky (1793-1856), ruso, estudió en la universidad de Kazan y de 1827 a 1846 fue profesor y rector de esa universidad. Presentó sus puntos de vista sobre los fundamentos de la geometría en un ensayo ante el departamento de matemáticas y física de la universidad en 1826. Sin embargo, el ensayo no fue publicado jamás y se perdió. Proporcionó su enfoque de la geometría no euclídea en una serie de ensayos, los dos primeros de los cuales fueron publicados en revistas en Kazan y el tercero en el *Journal für Mathematik*.[11] El primero fue titulado «Sobre los fundamentos de la geometría» y apareció en 1829-1830. El segundo, titulado «Nuevos fundamentos de la geometría con una teoría completa de las paralelas» (1835-1837), es una presentación mejor de las ideas de Lobatchevsky. Llamó a su nueva geometría, geometría imaginaria por razones que son tal vez ya visibles y que serán claras posteriormente. En 1840, publicó un libro en alemán *Geometrische Untersuchungen zur Theorie der Parallellinien (Investigaciones geométricas sobre la teoría de las paralelas)*.[12] En este libro lamenta el poco interés mostrado por sus escritos. A pesar de que perdió la vista, dictó una exposición completamente nueva de su geometría y la publicó en 1855 bajo el título de *Pangeometrie (Pangeometría)*.

Janos Bolyai (1802-1860), hijo de Wolfgang Bolyai, fue oficial del ejército húngaro. Escribió un ensayo de 26 páginas «La Ciencia del espacio absoluto» [13] sobre geometría no euclídea, que él llamó geometría absoluta. Fue publicado como un apéndice al trabajo de su padre *Tentamen Juventutem Studiosam in Elementa Matheseos* (Ensayo sobre los elementos de las matemáticas para jóvenes estudiosos). A pesar de que el libro (de dos volúmenes) apareció en 1832-1833, y, por tanto, después de una publicación de Lobatchevsky, Bolyai desarrolló aparentemente sus ideas sobre geometría no euclídea alrededor de 1825 y estaba convencido en esa época de que la nueva geometría no era contradictoria. En una carta a su

---

[11] *Jour. für Math.*, 17, 1837, 295-320.

[12] Las traducciones al inglés aparecen en Bonola. Véase la bibliografía al final de este capítulo.

[13] La traducción al inglés aparece en Bonola. Véase la bibliografía al final de este capítulo.

padre, fechada el 23 de noviembre de 1823, Janos escribió: «He hecho tan maravillosos descubrimientos que yo mismo me he perdido en el deslumbramiento.» El trabajo de Bolyai era tan parecido al de Lobatchevsky que cuando Bolyai vio por primera vez el trabajo de este último en 1835, pensó que había sido copiado de su propia publicación de 1832-1833. Por otro lado, Gauss leyó el ensayo de Janos Bolyai en 1832 y escribió a Wolfgang [14] que era incapaz de elogiarlo porque sería como elogiar su propio trabajo.

6. El contenido técnico de la geometría no euclídea

Gauss, Lobatchevsky y Bolyai se habían percatado de que el axioma euclídeo de la paralelas no podía ser demostrado tomando como base los otros nueve axiomas y que algún axioma adicional era necesario para fundamentar la geometría euclídea. Ya que el axioma de las paralelas era un hecho independiente, resultaba al menos lógicamente posible adoptar un enunciado contradictorio y desarrollar las consecuencias de ese nuevo conjunto de axiomas.

Para estudiar el contenido técnico de lo que estos hombres crearon, sólo es necesario tomar el trabajo de Lobatchevsky, ya que los tres hicieron casi la misma cosa. Lobatchevsky proporcionó, como sabemos, diversas versiones que difieren únicamente en pequeños detalles. Nuestro resumen estará basado en su ensayo de 1835-1837. Puesto que en los *Elementos* de Euclides se demuestran muchos teoremas que no dependen en absoluto del axioma de las paralelas, tales teoremas son válidos en la nueva geometría. Lobatchevsky dedica los primeros seis capítulos de su ensayo a la demostración de estos teoremas básicos. Desde un principio supone que el espacio es infinito, y así es capaz de demostrar que dos rectas no pueden intersecarse en más de un punto y que dos perpendiculares a la misma recta no se cortan.

En el séptimo capítulo, Lobatchevsky rechaza audazmente el postulado euclídeo de las paralelas y hace la siguiente suposición: dada una recta AB y un punto C (fig. 36.5), entonces todas las rectas que pasan por C caen dentro de dos clases con respecto a AB, a saber, la clase de las rectas que cortan AB y la clase de las que no lo hacen. A la última pertenecen las dos rectas p y q que forman la frontera

[14] *Werke*, 8, 220-221.

entre las dos clases. Estas dos líneas frontera son llamadas las rectas paralelas. Más precisamente, si *C* es un punto a una distancia perpendicular *a* de la recta *AB*, entonces existe un ángulo [15] $\pi(a)$ tal que todas las rectas que pasan por *C* y que forman con la perpendicular *CD* un ángulo menor que $\pi(a)$ intersecarán *AB;* todas las otras rectas que pasan por *C* no intersecan *AB.*[16] Las dos rectas que forman el ángulo $\pi(a)$ con *AB* son las paralelas y $\pi(a)$ es llamado el ángulo de paralelismo. Las otras rectas que no son paralelas y que pasan por *C* y las que no cortan a *AB* son llamadas rectas que no intersecan, aunque en el sentido de Euclides éstas son paralelas a *AB* y así en este sentido la geometría de Lobatchevsky contiene un número infinito de paralelas que pasan por *C.*

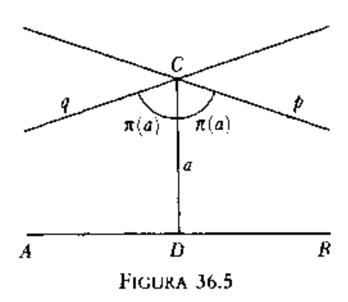


FIGURA 36.5

Si $\pi(a) = \pi/2$ entonces resulta el axioma euclídeo. Si no, entonces se sigue que $\pi(a)$ se incrementa y se aproxima a $\pi/2$ cuando *a* decrece a cero y $\pi(a)$ decrece y se aproxima a cero mientras *a* se hace infinito. La suma de los ángulos de un triángulo siempre es menor que π, decrece cuando se incrementa el área del triángulo y se aproxima a π cuando el área se aproxima a cero. Si los dos triángulos son semejantes, entonces son congruentes.

Ahora, Lobatchevsky se vuelve hacia la parte trigonométrica de su geometría. El primer paso es la determinación de $\pi(a)$. El resultado, si un ángulo central completo es 2π, es [17]

---

[15] El símbolo $\pi(a)$ es común y así es usado aquí. De hecho, el π de $\pi(a)$ no tiene nada que ver con el número π.

[16] La idea de que un ángulo específico pueda estar asociado con una longitud se debe a Lambert.

[17] Esta es una formulación especial. En su trabajo de 1840, Lobatchevsky pro-

$$\operatorname{tg}\frac{\pi(x)}{2} = e^{-x}, \tag{1}$$

de lo que se sigue que $\pi(0) = \pi/2$ y $\pi(+\infty) = 0$. La relación (1) es significativa en cuanto que a cada longitud x asocia un ángulo definido $\pi(x)$. Cuando $x = 1$, $tg\ [\pi(1)/2] = e^{-1}$, de tal forma que $\pi(1) = 40°24'$. Así la unidad de longitud es aquella longitud cuyo ángulo de paralelismo es $40°24'$. Esta unidad de longitud no tiene un significado físico directo. Físicamente puede ser una pulgada o una milla. Uno seleccionaría una interpretación física que haría la geometría físicamente aplicable.[18]

Después, Lobatchevsky deduce fórmulas conectando lados y ángulos de los triángulos planos de su geometría. En un ensayo de 1834, había definido el *cos x* y *sen x* para x real como las partes real e imaginaria de e^{ix}. La meta de Lobatchevsky consistía en proporcionar un fundamento puramente analítico para la trigonometría y hacerla independiente de la geometría euclídea. Las principales fórmulas trigonométricas de su geometría son (fig. 36.6).

$$cos\ \pi(a) = cos\ \pi(c) \operatorname{sen} A$$
$$sen\ A = cos\ B\ sen\ \pi(b)$$
$$\operatorname{sen} \pi(c) = sen\ \pi(a)\ sen\ \pi(b).$$

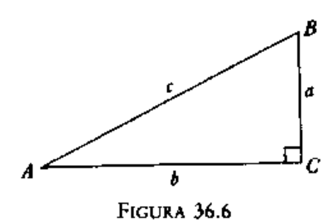


FIGURA 36.6

porciona lo que equivale a la forma comúnmente dada en los libros de texto y que también da Gauss, a saber,

$$\operatorname{tg}\frac{1}{2}\pi(x) = e^{-x/k} \tag{a}$$

donde k es una constante, llamada constante de espacio. Con fines teóricos, el valor de la k es indiferente. Bolyai también da la forma (a).

[18] En el caso de la relación $tg\ [\pi(x)/2] = e^{-x/k}$, la elección del valor de la x que debería corresponder a, digamos, $40°24'$, determinará la k.

Estas fórmulas son válidas en la trigonometría esférica ordinaria siempre que los lados tengan longitudes imaginarias. Esto es, si se reemplaza a, b y c en las fórmulas usuales de la trigonometría esférica por ia, ib e ic se obtienen las fórmulas de Lobatchevsky. Ya que las funciones trigonométricas de ángulos imaginarios son reemplazables por funciones hiperbólicas, se espera ver estas últimas funciones en las fórmulas de Lobatchevsky. Estas pueden ser obtenidas al usar la relación $tg\ (\pi(x)/2) = e^{-x/k}$. Así la primera de las fórmulas anteriores puede ser convertida en

$$\operatorname{senh}\frac{a}{k} = \operatorname{senh}\frac{c}{k}\operatorname{sen} A.$$

También, mientras que en la geometría esférica usual el área de un triángulo con los ángulos A, B y C es $r^2\ (A + B + C - \pi)$, en la geometría no euclídea ésta es $r^2\ [\pi - (A + B + C)]$, lo que equivale a reemplazar r por ir en la fórmula usual.

En su primer artículo de (1829-1830), Lovatchevsky derivó, trabajando con un triángulo infinitesimal, la fórmula

$$ds = \sqrt{(dy)^2 + \frac{(dx)^2}{\operatorname{sen}^2 \pi(x)}}$$

para el elemento de arco sobre la curva $y = f(x)$ en el punto (x,y). Así, puede ser calculada la circunferencia entera de un círculo de radio r, y es

$$C = \pi(e^r - e^{-r}).$$

La expresión para el área de un círculo resulta ser

$$A = \pi(e^{r/2} - e^{-r/2})^2.$$

También proporciona teoremas sobre el área de regiones planas y curvas y volúmenes de sólidos.

Las fórmulas de la geometría euclídea resultan de las fórmulas no euclídeas cuando las magnitudes son pequeñas. Así, si usamos el hecho que

$$e^r = 1 + r + \frac{r^2}{2!} + \ldots$$

y despreciamos para r pequeña todos los términos, excepto los dos primeros, por ejemplo,

$$C = \pi(e^r - e^{-r}) \sim \pi\{1 + r - (1 - r)\} = 2\pi r.$$

En el primer trabajo (1829-1830), Lobatchevsky también consideró la aplicabilidad de su geometría al espacio físico. La esencia de su argumento se apoya en el paralaje de las estrellas. Supóngase que E_1 y E_2 (fig. 36.7) son las posiciones de la tierra con seis meses de diferencia y S es una estrella. El paralaje p de S es la diferencia en las direcciones de E_1S y E_2S medida, digamos, a partir de la perpendicular E_1S'. Si E_1R es la paralela euclídea a E_2S, entonces, ya que E_1SE_2 es un triángulo isósceles, $\pi/2 - \sphericalangle SE_1E_2$ es la mitad del cambio en la dirección de la estrella, esto es, $p/2$. Este ángulo es 1″24 para la estrella Sirio (el valor de Lobatchevsky). En tanto ese ángulo no sea cero, la recta de E_1 a la estrella no puede ser paralela a TS ya que la línea corta a TS. Sin embargo, si existiera una cota inferior para los diversos paralajes de todas las estrellas, entonces la recta desde E_1 que forma un ángulo con E_1S' menor que esta cota inferior podía ser tomada como una paralela a TS por E_1 y esta geometría sería igualmente útil en lo que concierne a las medidas estelares. Pero, más adelante, Lobatchevsky demostró que la unidad de longitud en su geometría tendría que ser, físicamente, más de un millón de veces el radio de la trayectoria de la tierra. En otras palabras, la geometría de Lobatchevsky podía ser aplicable únicamente en un triángulo enormemente grande.

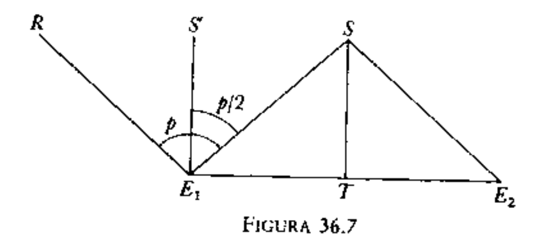


FIGURA 36.7

7. Las demandas de prioridad de Lobatchevsky y Bolyai

La creación de la geometría no euclídea es comúnmente usada como ejemplo de cómo una idea se les ocurre independientemente a diferentes personas casi al mismo tiempo. Algunas veces esto es visto como coincidencia pura, y algunas como prueba del espíritu del tiempo señalando una influencia en frentes ampliamente separados. La creación de la geometría no euclídea por Gauss, Lobatchevsky y Bolyai no es un ejemplo de creación simultánea, ni está bien justificado otorgarles su gran mérito a Lobatchevsky y Bolyai. Es cierto, como se mencionó anteriormente, que fueron los primeros en publicar abiertamente una geometría no euclídea y su acto mostró más coraje que el de Gauss. Sin embargo, la creación de la geometría no euclídea es difícilmente una contribución suya. Ya hemos señalado que aún Gauss fue precedido por Lambert, que Schweikart y Taurinus fueron creadores independientes, y que Schweikart y Taurinus publicaron su trabajo. Más aún, a Gauss se le debe el percatarse que la nueva geometría era aplicable al espacio físico.

Ambos, Lobatchevsky y Bolyai, le deben mucho a Gauss. Johann Martin Bartels (1769-1836), un buen amigo de Gauss, fue el maestro de Lobatchevsky en Kazan. De hecho, estuvieron juntos en Brunswick los años de 1805 a 1807. Posteriormente, Gauss y Bartels se mantuvieron en comunicación el uno con el otro. Es extremadamente difícil que Bartels no haya comunicado el progreso de la geometría no euclídea a Lobatchevsky, quien permaneció en Kazan como colega. En particular, Bartels debió haber conocido las dudas de Gauss en cuanto a la verdad de la geometría euclídea.

En cuanto a Janos Bolyai, su padre fue un amigo cercano de Gauss y compañero mientras estudiaban en Göttingen de 1796 a 1798. Wolfgang y Gauss no solamente continuaron en comunicación el uno con el otro, sino que discutieron específicamente el tema del axioma de las paralelas, como lo indica una de las citas anteriores. Wolfgang continuó trabajando duro sobre el problema del axioma de las paralelas y envió una pretendida demostración en 1804. Gauss le mostró que la prueba era falaz. Para 1817 Gauss tenía la certeza, no sólo de que el axioma no podía ser demostrado, sino de que una geometría no euclídea lógicamente consistente y físicamente aplicable podía ser construida. Además de su comunicación de 1799 a este efecto, Gauss transmitió sus últimos pensamientos libremente a Wolfgang. Wolfgang continuó trabajando sobre el problema hasta que

publicó su *Tentamen* de 1832-1833. Puesto que le recomendó a su hijo que estudiara el problema del axioma de las paralelas, con certeza debió haberle detallado todo lo que sabía.

Hay puntos de vista contradictorios. El matemático Friedrich Engel (1861-1941) pensaba que aunque Bartels, el maestro de Lobatchevsky, era amigo de Gauss, Lobatchevsky difícilmente podría haber sabido a través de su relación que Gauss había dudado de la corrección física del axioma de las paralelas. Pero este hecho, en sí mismo, fue crucial. Sin embargo, Engel dudó que Lobatchevsky hubiera sabido aún esto de Gauss, ya que Lobatchevsky había tratado desde 1816 de demostrar el axioma euclídeo de las paralelas; así, reconociendo la inutilidad de tales esfuerzos, finalmente creó en 1826 la nueva geometría. Janos Bolyai también intentó demostrar el axioma euclídeo de las paralelas hasta alrededor de 1820 y entonces se dirigió hacia la construcción de la nueva geometría. Pero estos esfuerzos continuos por demostrar el axioma de las paralelas no implican la ignorancia del pensamiento de Gauss.[19] Ya que nadie, ni aún Gauss, había demostrado que el axioma de Euclides de las paralelas no podía ser deducido a partir de los otros nueve axiomas, tanto Lobatchevsky como Bolyai pudieron haber decidido intentar resolver el problema. Después de fallar, apreciaron más rápidamente la sabiduría de los puntos de vista de Gauss en la materia.

En cuanto al contenido técnico contribuido por Lobatchevsky y Janos Bolyai, a pesar de que pudieron haber creado esto independientemente de sus predecesores y entre ellos mismos, los trabajos de Saccheri y Lambert, por no decir nada de los de Schweikart y Taurinus, eran ampliamente conocidos en Göttingen y les eran ciertamente conocidos a Bartels y Wolfgang Bolyai. Y Lobatchevsky admite por inferencia su conocimiento de los trabajos anteriores cuando se refiere en su ensayo de 1835-1837 a la futilidad de los esfuerzos durante más de dos mil años por resolver la cuestión del axioma de las paralelas.

---

[19] Sin embargo, véase George Bruce Halsted, *Amer. Math. Monthly*, 6, 1899, 166-172; y 7, 1900, 247-252.

8. Las implicaciones de la geometría no euclídea

Ya hemos señalado que la creación de la geometría no euclídea fue el paso de mayores consecuencias y más revolucionario desde los tiempos de los griegos. No trataremos todas las implicaciones de la materia en este momento. En su lugar, seguiremos el curso histórico de estos hechos. El impacto de la creación y la completa realización de su significación se postergó, ya que Gauss no publicó su trabajo sobre la materia y los trabajos de Lobatchevsky y Bolyai fueron ignorados durante aproximadamente treinta años. A pesar de que éstos eran conscientes de la importancia de su trabajo, los matemáticos exhibieron generalmente su reluctancia a abrigar ideas radicales. También, la materia clave en la geometría de los 1830 y 1840 fue la geometría proyectiva y por esta razón también el trabajo sobre la geometría no euclídea no atrajo a los matemáticos ingleses, franceses y alemanes. El tema atrajo la atención gracias a que el nombre de Gauss proporcionó peso a las ideas cuando las notas de Gauss y su correspondencia sobre geometría no euclídea fueron publicadas después de su muerte en 1855. Richard Baltzer (1818-1887), en un libro de 1866-1867, citó los trabajos de Lobatchevsky y Bolyai. Finalmente, desarrollos posteriores atrajeron a los matemáticos a la comprensión de la completa importancia de la geometría no euclídea.

Gauss sí vio las implicaciones más revolucionarias. El primer paso en la creación de la geometría no euclídea consistió en darse cuenta de que el axioma de las paralelas no podía ser demostrado sobre la base de los otros nueve axiomas. Era una aserción independiente, y así era posible adoptar un axioma contradictorio y desarrollar una geometría enteramente nueva. Esto lo hicieron Gauss y otros. Pero Gauss, habiéndose percatado de que la geometría euclídea no es necesariamente la geometría del espacio físico, esto es, no es necesariamente cierta, puso a la geometría en la misma clase que la mecánica y aseguraba que la cualidad de su certeza debía ser restringida a la de la aritmética (y su desarrollo en el análisis). Esta confianza en la aritmética es en sí misma curiosa. La aritmética, en esta época, no poseía un mínimo fundamento lógico. La seguridad de que la aritmética, el álgebra y el análisis ofrecían verdades acerca del mundo real provenía enteramente de su apoyo en la experiencia.

La historia de la geometría no euclídea revela, de una manera sorprendente, que los matemáticos son influenciados, no tanto por el razonamiento que ellos producen, sino más bien por el espíritu

de los tiempos. Saccheri había rechazado los teoremas extraños de la geometría no euclídea y concluido que Euclides había sido reivindicado. Pero cien años más tarde, Gauss, Lobatchevsky y Bolyai aceptaron confiadamente la nueva geometría. Creían que su nueva geometría era lógicamente consistente y de aquí que la geometría fuera tan válida como la de Euclides. Pero carecían de una prueba de su consistencia. A pesar de que demostraron muchos teoremas y no obtuvieron contradicciones evidentes, permanecía abierta la posibilidad de que se pudiera derivar una contradicción. De suceder esto, entonces la suposición de su axioma de las paralelas sería invalidada, como Saccheri había creído, y el axioma de las paralelas de Euclides sería una consecuencia de sus otros axiomas.

De hecho, Bolyai y Lobatchevsky consideraron la cuestión de la consistencia y estaban convencidos parcialmente de ella ya que su trigonometría era la misma para una esfera de radio imaginario y la esfera es parte de la geometría euclídea. Pero Bolyai no estaba satisfecho con esta evidencia ya que la trigonometría no es en sí misma un sistema matemático completo. Así, a pesar de la ausencia de cualquier demostración de consistencia, o de la aplicabilidad de la nueva geometría, que al menos habría servido como un elemento convincente, Gauss, Bolyai y Lobatchevsky aceptaron lo que sus predecesores habían considerado como absurdo. Esta aceptación fue un acto de fe. La cuestión de la consistencia de la geometría no euclídea permaneció como problema abierto otros cuarenta años.

Un punto más requiere atención y énfasis acerca de la creación de la geometría no euclídea. Existe una creencia popular de que Gauss, Bolyai y Lobatchevsky se encerraron en una esquina, jugaron a cambiar los axiomas de la geometría euclídea únicamente para satisfacer su curiosidad intelectual y así crearon una nueva geometría. Y ya que esta creación ha mostrado ser enormemente importante para la ciencia —una forma de geometría no euclídea que tenemos todavía que examinar ha sido usada en la teoría de la relatividad— muchos matemáticos han asentido a que la curiosidad puramente intelectual es una justificación suficiente para la exploración de cualquier idea matemática y que los valores de la ciencia seguirán casi seguramente como ha sucedido, a propósito, en el caso de la geometría no euclídea. Pero la historia de la geometría no euclídea no apoya esta tesis. Hemos visto que la geometría no euclídea llegó después de siglos de trabajo sobre el axioma de las paralelas. La preocupación por este axioma surgió del hecho de que debería ser,

como axioma, una verdad autoevidente. Ya que los axiomas de la geometría son nuestros hechos básicos acerca del espacio físico y grandes ramas de las matemáticas y de las ciencias físicas usan las propiedades de la geometría euclídea, los matemáticos desearon estar seguros de que estaban apoyándose sobre verdades. En otras palabras, el problema del axioma de las paralelas no era únicamente un problema físico genuino, sino tan fundamental como puede serlo un problema físico.

Bibliografía

Bonola, Roberto: *Non-Euclidean Geometry*, Dover (reimpresión), 1955.

Dunnington, G. W.: *Carl Friedrich Gauss*, Stechert-Hafner, 1960.

Engel, F., y Staeckel, P.: *Die Theorie der Parallellinien von Euklid bis auf Gauss*, 2 vols., B. G. Teubner, 1895.

— *Urkunden zur Geschichte der nichteuklidischen Geometrie*, B. G. Teubner, 1899-1913, 2 vols. El primer volumen contiene la traducción del ruso al alemán de los ensayos de Lobatchevsky de 1829-1830 y 1835-1837. El segundo es sobre el trabajo de los dos Bolyai.

Enriques, F.: «Prinzipien der Geometrie». *Encyk. der Math. Wiss.*, B. G. Teubner, 1907-1910, III AB1, 1-129.

Gauss, Carl F.: *Werke*. B. G. Teubner, 1900 y 1903, vol. 8, 157-268; vol. 9, 297-458.

Heath, Thomas L.: *Euclid's Elements*, Dover (reimpresión), 1956, vol. 1, 202-220.

Kagan, V.: *Lobatchevsky and his contribution to science*, Foreign Language Pub. House, Moscú, 1957. Hay versión castellana, *Lobachewki*, Moscú, Ed. Mir, 1986.

Lambert, J. H.: *Opera Mathematica*, 2 vols. Orell Fussli, 1946-1948.

Pasch, Moritz, y Max Dehn: *Vorlesungen uber neuere Geometrie*, 2.ª ed., Julius Springer, 1926, 185-238. Hay versión castellana, M. Pasch, *Lecciones de geometría moderna*, Madrid, Junta para Ampliación de Estudios, 1912.

Saccheri, Gerolamo: *Euclides ab Omni Naevo Vindicatus*, Trad. inglesa por G. B. Halsted en *Amer. Math. Monthly*, vols. 1-5, 1894-1898; también en Open Court Pub. Co., 1920 y Chelsea (reimpresión), 1970.

Schmidt, Franz, y Paul Staeckel: *Briefwechsel zwischen Carl Friedrich Gauss und Wolfgang Bolyai*, B. G. Teubner, 1899; Georg Olms (reimpresión), 1970.

Smith, Davis F.: *A source book in mathematics*, Dover (reimpresión), 1959, vol. 2, 351-388.

Sommerville, D. M. Y.: *The elements of non-euclidean geometry*, Dover (reimpresión), 1958.
Staeckel, P.: «Gauss als Geometer», *Nachrichten König. Ges. der Wiss. zu Gött.*, 1917, Beiheft, 25-142. También en Gauss: Werke, X_2. Von Walterhausen, W. Sartorius: *Carl Friedrich Gauss*, S. Hirzel, 1856; Springer Verlag (reimpresión), 1965.
Zacharias, M.: «Elementargeometrie und elementare nicht-euklidische Geometrie in synthetischer Behandlung». *Encyk. der Math. Wiss.*, B. G. Teubner, 1914-1931, III AB9, 859-1172.

Capítulo 37

LA GEOMETRIA DIFERENCIAL DE GAUSS Y RIEMANN

Tú, naturaleza, sé mi diosa; a tus leyes mis servicios están atados...

CARL F. GAUSS

1. Introducción

Debemos ahora recoger los hilos del desarrollo de la geometría diferencial, particularmente la teoría de superficies tal como fue fundada por Euler y extendida por Monge. Gauss dio el gran paso siguiente en la disciplina.

Empezando en 1816, Gauss dedicó una gran cantidad de trabajo a la geodesia y la cartografía. Su participación en mediciones físicas reales, sobre las que publicó muchísimos ensayos, estimuló su interés en la geometría diferencial y lo condujo a su ensayo definitivo de 1827 «Disquisitiones Generales circa Superficies curvas» («Investigaciones generales sobre superficies curvas»).[1] Sin embargo, Gauss propuso el concepto completamente nuevo de que una superficie es un espacio en ella misma, además de contribuir al tratamiento definitivo de la geometría diferencial de superficies situadas en un espacio tridimensional. Fue este concepto el que generalizó Riemann, abriendo desde entonces nuevas perspectivas en la geometría no euclídea.

---

[1] *Comm. Soc. Gött.*, 6, 1828, 99-146 = *Werke*, 4, 217-258.

2. La geometría diferencial de Gauss

Euler ya había introducido la idea (cap. 23, sec. 7) de que las coordenadas (x,y,z) de cualquier punto sobre una superficie pueden ser representadas en términos de dos parámetros u y v; esto es, las ecuaciones de la superficie están dadas por

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v). \tag{1}$$

Gauss partió de usar la representación paramétrica para el estudio sistemático de las superficies. A partir de estas ecuaciones paramétricas tenemos

$$dx = a\,du + a'\,dv, \quad dy = b\,du + b'\,dv, \quad dz = c\,du + c'\,dv \tag{2}$$

donde $a = x_u$, $a' = x_v$, y así en adelante. Por conveniencia, Gauss introduce los determinantes

$$A = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, \qquad B = \begin{vmatrix} c & a \\ c' & a' \end{vmatrix}, \qquad C = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$$

y la cantidad

$$\Delta = \sqrt{A^2 + B^2 + C^2}$$

que él supone no es idénticamente igual a cero.

La cantidad fundamental sobre cualquier superficie es el elemento de longitud de arco, que en coordenadas (x,y,z) es

$$ds^2 = dx^2 + dy^2 + dz^2. \tag{3}$$

Gauss usa las ecuaciones (2) para escribir (3) como

$$ds^2 = E(u, v)du^2 + 2F(u, v)du\,dv + G(u, v)dv^2, \tag{4}$$

donde

$$E = a^2 + b^2 + c^2, \quad F = aa' + bb' + cc', \quad G = a'^2 + b'^2 + c'^2.$$

Otra cantidad fundamental es el ángulo entre dos curvas sobre

una superficie. Una curva sobre una superficie está determinada por una cierta relación entre u y v, ya que entonces x, y y z se convierten en funciones de un parámetro, u o v, y las ecuaciones (1) se convierten en la representación paramétrica de una curva. Se dice en el lenguaje de las diferenciales que en un punto (u,v) una curva o la dirección de una curva emanando de un punto está dada por la proporción $du:dv$. Si entonces tenemos dos curvas o dos direcciones partiendo de (u,v), una dada por $du:dv$ y la otra por $du':dv'$, y si θ es el ángulo entre estas direcciones, Gauss muestra que

$$\cos\theta=\frac{E\,du\,du'+F(du\,dv'+du'\,dv)+G\,dv\,dv'}{\sqrt{E\,du^2+2F\,du\,dv+G\,dv^2}\sqrt{E\,du'^2+2F\,du'\,dv'+G\,dv'^2}}. \qquad (5)$$

Gauss se dedica luego al estudio de la curvatura de una superficie. Su definición de curvatura es una generalización a las superficies de la indicatriz usada para las curvas en el espacio por Euler, y usada para las superficies por Olinde Rodrigues.[2] En cada punto (x,y,z) sobre una superficie hay una normal cuya dirección está fijada. Gauss considera una esfera unitaria y selecciona un radio teniendo la dirección de la normal dirigida sobre la superficie. La elección del radio determina un punto (X,Y,Z) sobre la esfera. Si consideramos sobre la superficie cualquier región pequeña rodeando a (x,y,z), entonces existe una región correspondiente sobre la esfera rodeando a (X,Y,Z). La curvatura de la superficie en (x,y,z) está definida como el límite del radio del área de la región sobre la esfera al área de la región correspondiente sobre la superficie cuando las dos áreas tienden a sus puntos respectivos. Gauss evalúa esta proporción notando primero que el plano tangente en (X,Y,Z) sobre la esfera es paralelo al correspondiente a (x,y,z) sobre la superficie. De aquí que la proporción de las dos áreas sea la proporción de sus proyecciones sobre los planos tangentes respectivos. Para encontrar esta última proporción, Gauss realiza un número increíble de diferenciaciones y obtiene un resultado que es aún básico, a saber, que la curvatura (total) K de la superficie es

$$K=\frac{LN-M^2}{EG-F^2} \qquad (6)$$

[2] *Corresp. sur l'Ecole Poly.*, 3, 1814-1816, 162-182.

donde

$$L=\begin{vmatrix} x_{uu} & y_{uu} & z_{uu} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix}, \quad M=\begin{vmatrix} x_{uv} & y_{uv} & z_{uv} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix}, \quad N=\begin{vmatrix} x_{vv} & y_{vv} & z_{vv} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix}. \tag{7}$$

Entonces, Gauss demuestra que esta K es el producto de las dos curvaturas principales en (x,y,z), ya introducidas por Euler. En 1831, Sophie Germain[3] introdujo la noción de curvatura media, a saber, el promedio de las dos curvaturas principales.

Ahora Gauss hace una observación extremadamente importante. Cuando la superficie está dada mediante las ecuaciones paramétricas (1), las propiedades de la superficie parecen depender de las funciones x, y y z. Fijando u, digamos $u = 3$, y dejando variar v, se obtiene una curva sobre la superficie. Se obtiene una familia de curvas para los diversos valores fijos de u. De manera análoga, fijando v se obtiene una familia de curvas. Estas dos familias son las curvas paramétricas sobre la superficie, de tal forma que cada punto está dado por un par de números (c,d), digamos, donde $u = c$ y $v = d$ son las curvas paramétricas que pasan por un punto. Estas coordenadas no denotan necesariamente distancias más de lo que lo hacen las latitudes y las longitudes. Pensemos en una superficie en la cual las curvas paramétricas han sido determinadas de alguna manera. Entonces, afirma Gauss, las propiedades geométricas de la superficie están determinadas sólo por las E, F y G en la expresión (4) para ds^2. Estas funciones de u y v son todo lo que importa.

Sucede ciertamente, como es evidente a partir de (4) y (5), que las distancias y los ángulos sobre la superficie están determinados por las E, F y G. Pero la expresión fundamental de Gauss para la curvatura, la (6) arriba, depende de las cantidades adicionales L, M y N. Gauss ahora demuestra que

$$K=\frac{1}{2H}\left\{\frac{\partial}{\partial u}\left[\frac{F}{EH}\frac{\partial E}{\partial v}-\frac{1}{H}\frac{\partial G}{\partial u}\right]+\frac{\partial}{\partial v}\left[\frac{2}{H}\frac{\partial F}{\partial u}-\frac{1}{H}\frac{\partial E}{\partial v}-\frac{F}{EH}\frac{\partial E}{\partial u}\right]\right\} \tag{8}$$

donde $H = \sqrt{EG - F^2}$ y es igual a la Δ de Gauss definida ante-

[3] *Jour. für Math.*, 7, 1831, 1-29.

riormente. La ecuación (8), llamada la ecuación característica de Gauss, muestra que la curvatura K, y, en particular, a la vista de (6), la cantidad $LN - M^2$ depende únicamente de E, F y G. Ya que E, F y G son funciones sólo de las coordenadas paramétricas sobre la superficie, la curvatura también es sólo función de los parámetros y no depende en absoluto de sí o cómo yace la superficie en el espacio tridimensional.

Gauss observó que las propiedades de una superficie dependen únicamente de E, F y G. Sin embargo, muchas otras propiedades diferentes de la curvatura conciernen a las cantidades L, M y N, y no en la combinación $LN - M^2$ que aparece en la ecuación (6). Gaspare Mainardi (1800-1879),[4] independientemente de Delfino Codazzi (1824-1875),[5] realizó el trabajo analítico de la observación de Gauss y ambos proporcionaron dos relaciones adicionales en la forma de ecuaciones diferenciales que, junto con la ecuación característica de Gauss, teniendo k el valor en (6), determinan L, M y N en términos de E, F y G.

Más adelante, en 1867,[6] Ossian Bonnet (1819-1892) demostró el teorema de que cuando seis funciones satisfacen la ecuación característica de Gauss y las dos ecuaciones de Mainardi-Codazzi, entonces determinan una superficie única, con excepción de su posición y orientación en el espacio. Específicamente, si E, F y G y L, M y N están dadas como funciones de u y v que satisfacen la ecuación característica y las ecuaciones de Mainardi-Codazzi y si $EG - F \neq 0$, entonces existe una superficie dada por las tres funciones de x, y y z como funciones de u y v que tiene la primera forma fundamental

$$E\,du^2 + 2F\,du\,dv + G\,dv^2,$$

y L, M y N están relacionadas con E, F y G a través de (7). Esta superficie está determinada unívocamente, excepto por la posición en el espacio. (Para superficies reales con coordenadas reales (u,v), debemos tener $EG - F^2 > 0$, $E > 0$ y $G \geqslant 0$). El teorema de Bonnet es el análogo del teorema correspondiente sobre curvas (cap. 23, sec. 6).

El hecho de que las propiedades de una superficie dependen sólo

---

[4] *Giornale dell'Istituto Lombardo*, 9, 1856, 385-398.
[5] *Annali di Mat.* (3), 2, 1868-1869, 101-119.
[6] *Jour. de l'Ecole Poly.*, 25, 1867, 31-151.

de E, F y G tiene muchas implicaciones, algunas de las cuales dedujo Gauss en su ensayo de 1827. Por ejemplo, si una superficie es doblada sin estirarla o contraerla, las líneas coordenadas u = const. y v = const. permanecerán y también ds será la misma. De aquí que todas las propiedades de la superficie, en particular la curvatura, serán las mismas. Más aún, si dos superficies pueden ser puestas en una correspondencia uno-a-uno una con la otra, esto es, si $u' = \theta(u,v)$, $v' = \psi(u,v)$, donde u' y v' son las coordenadas de los puntos sobre la segunda superficie, y si el elemento de distancia en puntos correspondientes es el mismo sobre las dos superficies, esto es, si

$$E\,du^2 + 2F\,du\,dv + G\,dv^2 = E'\,du'^2 + 2F'\,du'\,dv' + G'\,dv'^2$$

donde las E, F y G son funciones de u y v y E', G' y G' son funciones de u' y v', entonces las dos superficies, que se dice son *isométricas*, deben tener la misma geometría. En particular, como señaló Gauss, deben tener la misma curvatura total en puntos correspondientes. A este resultado lo llamó Gauss un «teorema egregium», un teorema excelentísimo.

Como corolario, se sigue que para mover una parte de una superficie sobre otra parte (lo que significa preservar la distancia) una condición necesaria es que la superficie tenga curvatura constante. De esta manera, una parte de una esfera puede ser movida sin distorsión sobre otra, pero esto no puede ser realizado con un elipsoide. (Sin embargo, sí puede tener lugar un plegamiento al acomodar una superficie o parte de una superficie dentro de otra bajo una aplicación isométrica). A pesar de que las curvaturas en puntos correspondientes son iguales, si dos superficies no tienen curvatura constante, no están relacionadas necesariamente isométricamente. En 1839,[7] Ferdinand Minding (1806-1885) demostró que si dos superficies tienen curvatura constante e igual entonces una puede ser aplicada isométricamente sobre la otra.

Otra cuestión de gran importancia que Gauss estudió en su ensayo de 1827 es la de encontrar geodésicas sobre las superficies. (Liouville introdujo, en 1850, el término geodésica, que fue tomado de la geodesia). Este problema requiere del cálculo de variaciones, que Gauss utiliza. Enfoca este tema trabajando con la representación

[7] *Jour. für Math.*, 19, 1839, 370-387.

x, y, z y demuestra el teorema establecido por Jean Bernoulli de que la normal principal a una curva geodésica es normal a la superficie. (Así, la normal principal a un punto de un círculo de latitud sobre una esfera yace en el plano del círculo y no es normal a la esfera, mientras que la normal principal en un punto sobre el círculo de longitud es normal a la esfera). Cualquier relación entre u y v determina una curva sobre una superficie y la relación que proporciona una geodésica está determinada por una ecuación diferencial. Esta ecuación, de la que Gauss únicamente dice que es una ecuación de segundo orden en u y v, pero que no proporciona explícitamente, se escribe de muchas formas. Una es

$$\frac{d^2v}{du^2} = n\left(\frac{dv}{du}\right)^3 + (2m - \nu)\left(\frac{dv}{du}\right)^2 + (l - 2\mu)\frac{dv}{du} - \lambda, \tag{9}$$

donde n, m, μ, ν, l y λ son funciones de E, F y G.

Se debe ser cuidadoso al suponer la existencia de una geodésica única entre dos puntos sobre una superficie. Dos puntos cercanos sobre una superficie esférica tienen una geodésica única que los une, pero dos puntos diametralmente opuestos están unidos por una infinidad de geodésicas. Análogamente, dos puntos sobre el mismo generador de un cilindro circular están unidos mediante una geodésica a lo largo de la generatriz, pero también por un número infinito de hélices geodésicas. Si existe un solo arco geodésico entre dos puntos de una región, el arco proporciona la trayectoria mínima entre ellos en la región. Muchos matemáticos se dedicaron al problema de determinar de hecho las geodésicas sobre superficies particulares.

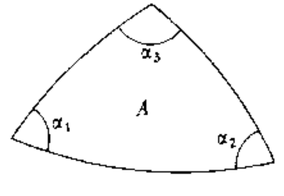


FIGURA 37.1

En su artículo de 1827, Gauss demostró un famoso teorema sobre curvatura para un triángulo formado por geodésicas (fig. 37.1). Sea K la curvatura variable de una superficie. $\iint K\, dA$ es entonces

la integral de esta curvatura sobre el área A. El teorema de Gauss, cuando se aplica al triángulo, establece que

$$\iint_A K\,dA = \alpha_1 + \alpha_2 + \alpha_3 - \pi;$$

esto es, la integral de la curvatura sobre un triángulo geodésico es igual al exceso de la suma de los ángulos sobre 180° o, cuando el ángulo suma es menor que π, al defecto con 180°. Este teorema, dice Gauss, debe ser considerado como un teorema elegantísimo. El resultado generaliza el teorema de Lambert (cap. 36, sec. 4), que afirma que el área de un triángulo esférico iguala el producto de su exceso esférico y al cuadrado del radio, ya que en un triángulo esférico K es constante e igual a $1/R^2$.

Debemos mencionar un trabajo también importante de la geometría diferencial de Gauss. Lagrange (cap. 23, sec. 8) había tratado la aplicación conforme de una superficie de revolución en un plano. En 1822, Gauss ganó un premio, propuesto por la Real Sociedad Danesa de Ciencias a un ensayo sobre el problema de encontrar la condición analítica para transformar cualquier superficie conformemente sobre cualquier otra superficie.[8] Su condición, que se refiere a la vecindad de puntos correspondientes sobre las dos superficies, es igual al hecho de que una función $P + iQ$, que no especificaremos con más detalle, de los parámetros T y U por medio de los cuales una superficie está representada es una función f de $p + iq$, que es la función correspondiente de los parámetros t y u por la cual está representada la otra superficie y $P - iQ$ es $f'(p - iq)$, donde f' es f o es obtenida de f reemplazando i por $-i$. La función f depende de la correspondencia entre las dos superficies, estando la correspondencia especificada por $T = T(t,u)$ y $U(t,u)$. Gauss no respondió a la cuestión de si y en qué manera una porción *finita* de la superficie puede ser aplicada conformemente sobre la otra superficie. Riemann trató este problema en su trabajo sobre funciones complejas (cap. 27, sec. 10).

El trabajo de Gauss en geometría diferencial es un punto de referencia en sí mismo. Pero sus implicaciones fueron mucho más profundas que lo que él mismo apreció. Hasta sus trabajos, las superficies habían sido estudiadas como figuras en un espacio euclídeo

---

[8] *Werke*, 4, 189-216.

tridimensional. Pero Gauss demostró que la geometría de una superficie podía ser estudiada concentrándose en la misma superficie. Si se introducen las coordenadas u y v que provienen de la representación paramétrica

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)$$

de la superficie en el espacio tridimensional y se usan las E, F y G determinadas de este modo, entonces se obtienen las propiedades euclídeas de esa superficie. Sin embargo, dadas estas coordenadas u y v sobre la superficie y la expresión para ds^2 en términos de E, F y G como funciones de u y v, todas las propiedades de la superficie se siguen de esta expresión. Esto sugiere dos ideas fundamentales. El primero es que la superficie puede ser considerada como un espacio en sí misma, ya que todas sus propiedades están determinadas por la ds^2. Se puede olvidar el hecho de que la superficie yace en un espacio tridimensional. ¿Qué clase de geometría posee la superficie si es considerada como un espacio en sí misma? Si uno toma como «líneas rectas» las geodésicas sobre esa superficie, entonces la geometría es no euclídea.

De esta manera, si se estudia la superficie de la esfera como un espacio en sí misma, tiene su propia geometría, y si además se usan las latitudes y longitudes familiares como coordenadas de los puntos, la geometría sobre esa superficie es no euclídea, ya que las «líneas rectas» o geodésicas son arcos de los círculos máximos sobre la superficie. Sin embargo, la geometría de la superficie esférica es euclídea si se la considera como una superficie en el espacio tridimensional. La distancia mínima entre dos puntos sobre la superficie es entonces el segmento lineal de una geometría euclídea tridimensional (aunque no yace sobre la superficie). El trabajo de Gauss implica que existen geometrías no euclídeas al menos sobre superficies consideradas como espacios en sí mismas. No está claro si Gauss vio o no esta interpretación de su geometría de superficies.

Se puede ir más allá. Se podría pensar que están determinadas las propias E, F y G para una superficie mediante las ecuaciones paramétricas (1). Pero se podría empezar con la superficie, introducir las dos familias de curvas paramétricas y entonces escoger casi arbitrariamente funciones E, F y G de u y v. Entonces la superficie tiene una geometría determinada por estas E, F y G. Esta geometría es intrínseca a la superficie y no tiene conexión con el espacio que la

rodea. Consecuentemente, la misma superficie puede tener *diferentes* geometrías, dependiendo de la elección de las funciones E, F y G.

Las implicaciones son más profundas. Si se pueden escoger diferentes conjuntos de *E, F* y *G* y, de esta manera, determinar diferentes geometrías sobre la misma superficie, ¿por qué no escoger diferentes funciones distancia en nuestro espacio tridimensional ordinario? La función de distancia común en coordenadas rectangulares es, por supuesto, $ds^2 = dx^2 + dy^2 + dz^2$ y esto es obligatorio si se empieza con la geometría euclídea, ya que es justamente el enunciado analítico del teorema pitagórico. Sin embargo, dadas las mismas coordenadas rectangulares para los puntos del espacio, se podría elegir una expresión diferente para ds^2 y obtener una geometría bastante diferente para ese espacio, una geometría no euclídea. Riemann tomó y desarrolló esta extensión a cualquier espacio de las ideas de Gauss que se obtuvo por primera vez al estudiar las superficies.

3. El enfoque de Riemann de la geometría

Las dudas acerca de lo que podemos creer de la geometría del espacio físico, originadas por el trabajo de Gauss, Lobatchevsky y Bolyai, estimularon una de las más grandes creaciones del siglo XIX: la geometría riemanniana. El creador fue Georg Bernhard Riemann, el filósofo más profundo de la geometría. Aunque los detalles de los trabajos de Lobatchevsky y Bolyai le eran desconocidos, eran conocidos de Gauss, y Riemann ciertamente conoció las dudas de Gauss en cuanto a la verdad y necesaria aplicabilidad de la geometría euclídea. De esta manera, en el campo de la geometría, Riemann siguió a Gauss mientras que en la teoría de funciones siguió a Cauchy y Abel. Las enseñanzas del psicólogo Johann Friedrich Herbart (1776-1841) influenciaron sus investigaciones de geometría.

Gauss propuso a Riemann el tema de los fundamentos de la geometría, sobre el que debería pronunciar su conferencia calificadora, la *Habilitationsvortag*, para el título de *Privatdozent*. Pronunció la conferencia de 1854, dirigida al cuerpo académico de Göttingen con Gauss presente, y fue publicada en 1868 con el título: «Über die Hypothese, welche der Geometrie zu Grunde liegen» («Sobre las hipótesis en que se funda la geometría»).[9]

---

[9] *Abh. der Ges. der Wiss. zu Gott.*, 13, 1868, 1-20 = *Werke*, 2.ª ed. 272-287. En

En un artículo sobre la conducción del calor, que Riemann escribió en 1861 para competir por un premio ofrecido por la Academia de Ciencias de París, y al que comúnmente se cita como el *Pariserarbeit*, Riemann encontró la necesidad de considerar más profundamente sus ideas sobre geometría, y aquí proporcionó algunas elaboraciones técnicas de su ensayo de 1854. Este ensayo de 1861, que no ganó el premio, se publicó póstumamente en 1876 en sus *Obras Completas*.[10] En la segunda edición de sus *Werke*, Heinrich Weber explica en una nota el material, muy resumido, de Riemann.

La geometría del espacio de Riemann no fue únicamente una extensión de la geometría diferencial de Gauss. Reconsideró el enfoque total del estudio del espacio. Riemann estudió la cuestión de en qué podemos estar seguros acerca del espacio físico. ¿Qué condiciones o hechos se presuponen en la propia experiencia del espacio antes de que nosotros determinemos por experiencia los axiomas particulares que se cumplen en el espacio físico? Uno de los objetivos de Riemann fue demostrar que los axiomas particulares de Euclides eran empíricos en lugar de, como se había creído, verdades autoevidentes. Adoptó el enfoque analítico, ya que en las demostraciones geométricas podemos ser desviados por nuestras percepciones para aceptar hechos no reconocidos explícitamente. Así, la idea de Riemann era apoyarse en el análisis para empezar con lo que seguramente es *a priori* acerca del espacio y deducir las consecuencias necesarias. Cualesquiera otras propiedades del espacio serían entonces tomadas como empíricas. Gauss se había preocupado de este mismo problema, pero sólo se publicó el ensayo sobre superficies curvas de esta investigación. La búsqueda de Riemann por lo que es *a priori* lo llevó a estudiar el comportamiento local del espacio, o, en otras palabras, el enfoque de la geometría diferencial como opuesto a la consideración del espacio como un todo como se encuentra en Euclides o en la geometría no euclídea de Gauss, Bolyai y Lobatchevsky. Antes de examinar los detalles, debemos estar advertidos de que las ideas de Riemann expresadas en su conferencia y en su manuscrito de 1854 son vagas. Una razón es que Riemann las adaptó para su audiencia, el profesorado de Göttingen al completo.

---

los *Collected Mathematical Papers* de W. K. Clifford se encuentra en traducción al inglés. También en *Nature*, 8, 1873, 14-36 y en D. E. Smith, *A source book in mathematics*, 411-425.

[10] *Werke*, 2.ª ed., 1892, 391-404.

Parte de la vaguedad surge a partir de las consideraciones filosóficas con las que Riemann inició su artículo.

Guiado en gran parte por la geometría intrínseca de Gauss de las superficies en el espacio euclídeo, Riemann desarrolló una geometría intrínseca para cualquier espacio. Prefirió tratar una geometría n-dimensional a pesar de que el caso tridimensional era claramente el importante, y habla del espacio n-dimensional como una variedad. Un punto en la variedad de n dimensiones está representado mediante la asignación de valores especiales a n parámetros variables, x_1, x_2, ..., x_n, y el agregado de todos esos tales posibles puntos constituye la variedad n-dimensional en sí misma, de la misma manera que el agregado de los puntos sobre la superficie constituye la propia superficie. Se llama a los n parámetros variables las coordenadas de la variedad. Cuando las x varían continuamente, los puntos recorren la variedad.

Puesto que Riemann creía que sólo conocemos el espacio localmente, empezó por definir la distancia entre dos puntos genéricos cuyas coordenadas correspondientes difieren sólo en cantidades infinitesimales. Supone que el cuadrado de la distancia es

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}\, dx_i\, dx_j, \tag{10}$$

donde las g_{ij} son funciones de las coordenadas x_1, x_2, ... x_n, $g_{ij} = g_{ji}$ y el lado derecho de (10) es siempre positivo para todos los valores posibles de las dx_i. Esta expresión para las ds^2 es una generalización de la fórmula de la distancia euclídea.

$$ds^2 = dx_1^2 + dx_2^2 + \ldots + dx_n^2.$$

Menciona la posibilidad de suponer que la ds^2 es la raíz cuarta de una función homogénea del cuarto grado en las diferenciales dx_1, dx_2, ..., dx_n, pero no siguió con esta posibilidad. Al permitir que las g_{ij} sean funciones de las coordenadas, Riemann prevé la posibilidad de que la naturaleza del espacio pudiera variar de un punto a otro.

Aunque en su ensayo de 1854 Riemann no estableció explícitamente las definiciones siguientes, indudablemente las tenía en su mente, porque se asemejan a lo que Gauss hizo para las superficies. Una curva sobre una variedad riemaniana está dada por el conjunto de n funciones.

$$x_1 = x_1(t), x_2 = x_2(t), \ldots, x_n = x_n(t). \qquad (11)$$

Entonces la longitud de la curva entre $t = \alpha$ y $t = \beta$ está definida por

$$l = \int_\alpha^\beta ds = \int_\alpha^\beta \frac{ds}{dt}\,dt = \int_\alpha^\beta \sqrt{\sum_{i,j=1}^{n} g_{ij}\,\frac{dx_i}{dt}\,\frac{dx_j}{dt}}\;dt. \qquad (12)$$

La curva más corta entre dos puntos dados, $t = \alpha$ y $t = \beta$, la geodésica, se determina entonces por el método del cálculo de variaciones. En la notación de esa materia es la curva para la cual $\delta \int_\alpha^\beta ds = 0$. Se debe determinar entonces las funciones particulares de la forma (11) que proporcionan esta trayectoria mínima entre los dos puntos. En términos del parámetro de longitud de arco, las ecuaciones de las geodésicas resultan ser

$$\frac{d^2x_i}{ds^2} + \sum_{\lambda,\mu} \left\{ {\lambda\mu \atop i} \right\} \frac{dx_\lambda}{ds}\,\frac{dx_\mu}{ds} = 0, \qquad i, \lambda, \mu = 1, 2, \ldots, n.$$

Este es un sistema de n ecuaciones diferenciales ordinarias de segundo orden.[11]

El ángulo θ entre dos curvas que se cortan en el punto $(x_1, x_2, \ldots, x_n)$, estando una curva determinada por las direcciones dx_i/ds, $i = 1, 2, \ldots, n$, y la otra por dx_i'/ds', $i = 1, 2, \ldots, n$, donde las primas indican valores que pertenecen a la segunda dirección, está definido por la fórmula

$$\cos\theta = \sum_{i,i'=1}^{n} g_{ii'}\,\frac{dx_i}{ds}\,\frac{dx_i'}{ds'}\,. \qquad (13)$$

Siguiendo los procedimientos que Gauss usó para las superficies, se desarrolla una geometría métrica n-dimensional sobre la base de las definiciones anteriores. Todas las propiedades métricas están determinadas por los coeficientes g_{ij} en la expresión para las ds^2.

La noción de curvatura de una variedad es el segundo concepto importante en el ensayo de Riemann de 1854. A través de ésta, Riemann buscó caracterizar el espacio euclídeo y, más generalmente,

[11] Para el significado del símbolo de la llave véase más adelante (19). Riemann no proporcionó estas ecuaciones explícitamente.

espacios en los que las figuras pueden ser movidas sin cambio en la forma o en la magnitud. La noción de Riemann de curvatura para cualquier variedad n-dimensional es una generalización de la noción de Gauss de curvatura total para las superficies. Al igual que la noción de Gauss, la curvatura de la variedad está definida en términos de cantidades determinables en la propia variedad y no hay necesidad de pensar en la variedad como si estuviera en alguna variedad de dimensión superior.

Dado un punto P en la variedad n-dimensional, Riemann considera una variedad bidimensional en el punto y en la variedad n-dimensional. Tal variedad bidimensional está formada por un único conjunto infinito de geodésicas que pasan por el punto y son tangentes a una sección plana de la variedad por el punto P. Ahora se puede describir una geodésica por medio del punto P y una dirección en ese punto. Sean $dx'_1, dx'_2, \ldots, dx'_n$ la dirección de una geodésica y $dx''_1, dx''_2, \ldots, dx''_n$ la dirección de otra. Entonces, la i-ésima dirección de cualquiera del conjunto infinito de geodésicas en P está dado mediante

$$dx_i = \lambda' dx'_i + \lambda'' dx''_i$$

(sujeta a la condición $\lambda'_2 + \lambda''_2 + 2\lambda'\lambda'' \cos\theta = 1$ que viene de la condición $\Sigma g_{ij}(dx_i/ds)(dx_j/ds) = 1$). Este conjunto de geodésicas forma una variedad bidimensional que tiene una curvatura gaussiana. Ya que existe una infinidad de tales variedades bidimensionales en P, obtenemos un número infinito de curvaturas en ese punto en la variedad n-dimensional. Pero de $n(n-1)/2$ de estas medidas de curvatura se deduce el resto. Se deriva una expresión explícita para la medida de la curvatura. Riemann hizo esto en su ensayo de 1861 y lo daremos más adelante. Para una variedad que es una superficie, la curvatura de Riemann es exactamente la curvatura total de Gauss. Hablando estrictamente, la curvatura de Riemann, como la de Gauss, es una propiedad de la métrica impuesta sobre la variedad y no de la misma variedad.

Después de que Riemann completase su investigación general de geometría n-dimensional y demostrase cómo se introduce la curvatura, consideró variedades más restringidas, en las cuales las formas espaciales finitas debían ser capaces de movimiento sin cambio de tamaño o forma y debían ser capaces de rotación en cualquier dirección. Esto lo condujo a espacios de curvatura constante.

Cuando todas las medidas de curvatura en un punto son las mismas e iguales a todas las medidas en cualquier otro punto, obtenemos lo que Riemann llama una variedad de curvatura constante. Sobre tal variedad es posible tratar figuras congruentes. En el artículo de 1854, Riemann proporcionó los siguientes resultados, pero no dio detalles: si α es la medida de curvatura, la fórmula para el elemento infinitesimal de distancia sobre una variedad de curvatura *constante* se convierte (en un sistema de coordenadas adecuado) en

$$ds^2 = \frac{\sum_{i=1}^{n} dx_i^2}{(1 + (\alpha/4) \Sigma x_i^2)^2} . \qquad (14)$$

Riemann pensó que la curvatura α debía ser positiva o cero, de tal forma que cuando $\alpha > 0$ obtenemos un espacio esférico y cuando $\alpha = 0$ obtenemos un espacio euclídeo, e inversamente. También creía que si un espacio estaba infinitamente extendido, la curvatura debía ser cero. Sugirió, sin embargo, que podía existir una superficie real de curvatura negativa constante.[12]

Para continuar con Riemann, para $\alpha = a^2 > 0$ y $n = 3$ obtenemos una geometría esférica tridimensional, aunque no podemos visualizarla. El espacio es finito en extensión pero sin límites. Todas las geodésicas en él son de longitud constante, a saber, $2\pi/a$, y regresan sobre sí mismas. El volumen del espacio es $2\pi^2/a^3$. Para $a^2 > 0$ y $n = 2$ obtenemos el espacio de la superficie esférica ordinaria. Las geodésicas son, por supuesto, los círculos máximos y son finitos. Aún más, dos cualesquiera se intersecan en dos puntos. De hecho, no está claro si Riemann consideró las geodésicas de una superficie de curvatura positiva constante como cortándose en uno o dos puntos. Posiblemente pretendía lo último. Félix Klein señaló más tarde (véase el próximo capítulo) que estaban envueltas dos geometrías distintas.

Riemann también señala una distinción, de la cual se hablará más tarde, entre la falta de límite e infinidad del espacio [como en el caso de la superficie de una esfera]. Lo ilimitado, dice, tiene una credi-

---

[12] Tales superficies ya le eran conocidas a Ferdinand Minding (*Jour. für Math.*, 19, 1839, 370-387, 378-380, en particular), incluyendo la más reciente llamada después pseudoesfera (véase cap. 38, sec. 2). Véase también: Gauss, *Werke*, 8, 265.

bilidad empírica mayor que cualquier otro hecho derivado empíricamente, tal como la extensión infinita.

Al final de este ensayo, Riemann nota que ya que el espacio físico es una clase especial de variedad, la geometría de ese espacio no puede ser derivada únicamente de las nociones generales de las variedades. Las propiedades que distinguen el espacio físico de cualesquiera otras variedades extendidas triplemente han de ser obtenidas únicamente por medio de la experiencia. Añade: «Queda por resolver la cuestión de conocer en qué medida y hasta qué punto estas hipótesis acerca de variedades son confirmadas por la experiencia.» En particular, los axiomas de la geometría euclídea pueden ser ciertos sólo aproximadamente del espacio físico. Como Lobatchevsky, Riemann pensaba que la astronomía decidiría qué geometría se ajusta al espacio. Finaliza su ensayo con esta observación profética: «O, por tanto, la realidad que subyace al espacio debe formar una variedad discreta o debemos buscar la base de sus relaciones métricas fuera de ella, en las fuerzas envolventes que actúan sobre ella... Esto nos conduce al dominio de otra ciencia, la de la física, a la que el objeto de nuestro trabajo no nos permite ir hoy.»

William Kingdon Clifford desarrolló este punto:[13]

Mantengo de hecho: 1) que las porciones pequeñas del espacio son de una naturaleza análoga a pequeñas colinas sobre una superficie que es en promedio plana. 2) Que esta propiedad de ser curvada o distorsionada es transmitida continuamente de una porción del espacio a otra a la manera de una onda. 3) Que esta variación de la curvatura del espacio es realmente lo que sucede en aquel fenómeno que nosotros llamamos el movimiento de la materia, ya sea ponderable o etéreo. 4) Que en este mundo físico nada tiene lugar excepto esta variación, sujeta, posiblemente, a la ley de continuidad.

Las leyes ordinarias de la geometría euclídea no son válidas para un espacio cuya curvatura cambia, no solamente con el lugar sino debido al movimiento de la materia, de un tiempo a otro. Añadió que una investigación más exacta de las leyes físicas no sería capaz de ignorar estas «colinas» en el espacio. De esta manera, Riemann y Clifford, contrariamente a la gran mayoría de los otros geómetras, sintieron la necesidad de asociar la materia con el espacio para poder

[13] *Proc. Camb. Phil. Soc.*, 2, 1870, 157-158 = *Math. Papers*, 20-22.

determinar lo que es cierto del espacio físico. Esta línea de pensamiento conduce, por supuesto, a la teoría de la relatividad.

En su *Pariserarbeit* (1861), Riemann regresó a la cuestión de cuándo un espacio riemanniano cuya métrica es

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}\, dx_i\, dx_j \tag{15}$$

puede ser un espacio de curvatura constante o incluso un espacio euclídeo. Sin embargo, formuló la cuestión más general de cuándo una métrica tal como (15) puede ser transformada por las ecuaciones

$$x_i = x_i(y_1, y_2, ..., y_n), \quad i = 1, 2, ..., n \tag{16}$$

en una métrica dada

$$ds'^2 = \sum_{i,j=1}^{n} h_{ij}\, dy_i\, dy_j \tag{17}$$

entendiéndose por supuesto, que ds igualará a ds', de tal forma que las geometrías de los dos espacios serían la misma excepto por la elección de coordenadas. La transformación (16) no siempre es posible porque, como señala Riemann, hay $n(n + 1)/2$ funciones independientes en (15), mientras que la transformación introduce únicamente n funciones que pueden ser usadas para convertir las g_{ij} en las h_{ij}.

Para discutir la cuestión general, Riemann introdujo cantidades especiales p_{ijk}, que debemos reemplazar por los más familiares símbolos de Christoffel entendiéndose que

$$p_{ijk} = \left[{jk \atop i}\right].$$

Los símbolos de Christoffel, denotados en varias formas, son

$$\Gamma_{\alpha\beta,\lambda} = \left[{\alpha\beta \atop \lambda}\right] = [\alpha\beta, \lambda] = \frac{1}{2}\left(\frac{\partial g_{\alpha\lambda}}{\partial x_\beta} + \frac{\partial g_{\beta\lambda}}{\partial x_\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x_\lambda}\right) \tag{18}$$

$$\Gamma^\lambda_{\alpha\beta} = \left\{{\alpha\beta \atop \lambda}\right\} = \{\alpha\beta, \lambda\} = \sum_i g^{i\lambda}\left[{\alpha\beta \atop i}\right] \tag{19}$$

donde $g^{i\lambda}$ es el cofactor dividido por g de $g_{i\lambda}$ en el determinante de

g. Riemann también introdujo lo que ahora es conocido como el símbolo riemanniano de cuatro índices

$$(\mu\lambda, jk) = R_{\lambda\mu,jk} = \frac{\partial\Gamma_{\lambda j,\mu}}{\partial x_k} - \frac{\partial\Gamma_{\lambda k,\mu}}{\partial x_j} + \sum_{i,\alpha} g^{i\alpha}(\Gamma_{\lambda k,\alpha}\Gamma_{\mu j,i} - \Gamma_{\lambda j,\alpha}\Gamma_{\mu k,i}). \tag{20}$$

Más adelante, Riemann muestra que una condición necesaria para que ds^2 sea transformable en ds'^2 es

$$(\alpha\delta, \beta\gamma)' = \sum_{r,k,i,h} (rk, ih) \frac{\partial x_r}{\partial y_\alpha} \frac{\partial x_i}{\partial y_\beta} \frac{\partial x_h}{\partial y_\gamma} \frac{\partial x_k}{\partial y_\delta} \tag{21}$$

donde la parte izquierda del símbolo se refiere a cantidades formadas para las métricas ds' y (21) se cumple para todos los valores de α, β, γ, δ, cada uno de los cuales varía de 1 a n.

Ahora Riemann se dirige a la cuestión específica de cuándo una ds^2 dada puede ser transformada en una con coeficientes constantes. Primero deriva una expresión explícita para la curvatura de una variedad. La definición general ya dada en el ensayo de 1854 hace uso de las líneas geodésicas partiendo de un punto O del espacio. Sean d y δ que determinan dos vectores o direcciones de geodésicas saliendo de O (cada dirección es especificada por las componentes de la tangente a la geodésica). Considérese entonces el haz de vectores geodésicos partiendo de O y dado por $kd + \lambda\delta$ donde k y λ son parámetros. Si se piensa en d y δ como operando sobre las $x_i = f_i(t)$ que describe cualquier otra curva, entonces existe un significado para la diferencial segunda $(kd + \lambda\delta)^2 = k^2d^2 + 2k\lambda d\delta + \lambda^2\delta^2$. Entonces Riemann forma

$$\Omega = \delta\delta \sum g_{ij}dx_idx_j - 2d\delta \sum g_{ij}dx_idx_j + dd \sum g_{ij}dx_idx_j. \tag{22}$$

Aquí se entiende que las d y δ operan formalmente sobre las expresiones que las siguen (y d y δ conmutan) de tal forma que

$$\delta\delta \sum g_{ij}dx_idx_j = \delta\left[\sum (\delta g_{ij})dx_idx_j + \sum g_{ij}((\delta dx_i)dx_j + dx_i\delta dx_j)\right] \tag{23}$$

y $\delta g_{ij} = \Sigma_r (\partial g_{ij}/\partial x_r)\, \delta x_r$. Si se calcula Ω se encuentra que todos los términos involucrando diferenciales terceras de una función se anu-

lan. Sólo los términos en que aparecen δx_i, dx_i, $\delta^2 x_i$, δdx_i y $d^2 x_i$ permanecen. Calculando estos términos y usando la notación

$$p_{ik} = dx_i \delta x_k - dx_k \delta x_i$$

Riemann obtiene

$$[\Omega] = \sum_{i,k,r,s} (ik, rs) p_{ik} p_{rs}. \quad (24)$$

Ahora hagamos

$$4\Delta^2 = \sum g_{ij} dx_i dx_j \cdot \sum g_{ij} \delta x_i \delta x_j - \left(\sum g_{ij} dx_i \delta x_j \right)^2.$$

Entonces la curvatura K de una variedad riemaniana es

$$K = -\frac{[\Omega]}{8\Delta^2}. \quad (25)$$

La conclusión global es que la condición necesaria y suficiente para que una ds^2 dada puede ser llevada a la forma (para $n = 3$)

$$ds'^2 = c_1\, dx_1^2 + c_2\, dx_2^2 + c_3\, dx_3^2, \quad (26)$$

donde las c_i son constantes, es que todos los símbolos $(\alpha\beta, \gamma\delta)$ sean cero. En el caso de que las c_i sean todas positivas, las ds' pueden ser reducidas a $dy_1^2 + dy_2^2 + dy_3^2$, esto es, el espacio es euclídeo. Como podemos ver por el valor de $[\Omega]$, cuando la K es cero, el espacio es esencialmente euclídeo.

Es útil notar que la curvatura de Riemann para una variedad n-dimensional se reduce a la curvatura total de Gauss de una superficie. De hecho cuando

$$ds^2 = g_{11}\, dx_1^2 + 2g_{12}\, dx_1 dx_2 + g_{22}\, dx_2^2,$$

de los dieciséis símbolos $(\alpha\beta, \gamma\delta)$, 12 son cero y para los cuatro restantes tenemos $(12,12) = -(12,21) = -(21,12) = (21,21)$. Así, la K de Riemann se reduce a

$$K = \frac{(12, 12)}{g}.$$

Usando (20) se puede demostrar que esta expresión es igual a la de Gauss para la curvatura total de una superficie.

4. Los sucesores de Riemann

El ensayo de Riemann de 1854 fue publicado en 1868, dos años después de su muerte; creó gran interés, y muchos matemáticos se apresuraron a completar y extender las ideas que él había bosquejado. Los sucesores inmediatos de Riemann fueron Beltrami, Christoffel y Lipschitz.

Eugenio Beltrami (1835-1900), profesor de matemáticas en Bolonia y otras universidades italianas, quien conocía el artículo de Riemann de 1854, pero que aparentemente no conocía el de 1861, abordó la cuestión de demostrar que la expresión general para ds^2 se reduce a la forma (14) dada por Riemann para un espacio de curvatura constante.[14] Además de este resultado y de haber demostrado otras cuantas aserciones de Riemann, Beltrami se dedicó al tema de los invariantes diferenciales, que consideraremos en la siguiente sección.

Elwin Bruno Christoffel (1829-1900), que fuera profesor de matemáticas en Zurich y más tarde en Estrasburgo, avanzó las ideas contenidas en ambos ensayos de Riemann. En dos ensayos clave,[15] la mayor preocupación de Christoffel consistió en reconsiderar y amplificar el tema tratado un tanto escuetamente por Riemann en su ensayo de 1861, a saber, cuándo una forma

$$F = \sum_{i,j} g_{ij}\, dx_i\, dx_j$$

puede ser transformada en otra

$$F' = \sum_{i,j} g'_{ij}\, dy_i\, dy_j.$$

---

[14] *Annali di Mat.*, (2), 2, 1868-1869, 232-255 = *Opera Mat.*, 1, 406-429.

[15] *Jour. für Math.*, 70, 1869, 46-70 y 241-245 = *Ges. Math. Abh.*, 1, 352 sgs., 378 sgs.

Christoffel buscó condiciones necesarias y suficientes. Incidentalmente, en este ensayo introdujo los símbolos de Christoffel.

Consideremos primero el caso bidimensional donde

$$F = a\,dx^2 + 2b\,dx\,dy + c\,dy^2$$

y

$$F' = A\,dX^2 + 2B\,sX\,dY + C\,dY^2$$

y supongamos que las x y las y se expresan como funciones de X y Y de tal forma que F se convierte en F' bajo la transformación. Por supuesto, $dx = (\partial x/\partial X)\,dX + (\partial x/\partial Y)\,dY$. Ahora, cuando x, y, dx y dy son reemplazadas en F por sus valores en X y Y y cuando se igualan los coeficientes en esta nueva forma de F con los de F', se obtiene

$$a\left(\frac{\partial x}{\partial X}\right)^2 + 2b\,\frac{\partial x}{\partial X}\frac{\partial y}{\partial X} + c\left(\frac{\partial y}{\partial X}\right)^2 = A$$

$$a\,\frac{\partial x}{\partial X}\frac{\partial x}{\partial Y} + b\left(\frac{\partial x}{\partial X}\frac{\partial y}{\partial Y} + \frac{\partial x}{\partial Y}\frac{\partial y}{\partial X}\right) + c\,\frac{\partial y}{\partial X}\frac{\partial y}{\partial Y} = B$$

$$a\left(\frac{\partial x}{\partial Y}\right)^2 + 2b\left(\frac{\partial x}{\partial Y}\frac{\partial y}{\partial Y}\right) + c\left(\frac{\partial y}{\partial Y}\right)^2 = C.$$

Estas son tres ecuaciones diferenciales para x e y como funciones de X e Y. Si se resuelven, entonces sabemos cómo transformar a F en F'. Sin embargo, hay sólo dos funciones relacionadas. Entonces deben existir algunas relaciones entre a, b y c, por un lado, y A, B y C por el otro. Diferenciando las tres ecuaciones anteriores y con pasos algebraicos posteriores se demuestra que la relación es $K = K'$.

Para el caso de n variables, Christoffel utiliza la misma técnica. Empieza con

$$F = \sum g_{rs}\,dx_r\,dx_s$$

y

$$F' = \sum g'_{rs}\,dy_r\,dy_s.$$

La transformación es

$$x_i = x_i(y_1, y_2, ..., y_n), \qquad i = 1, 2, ..., n$$

y hace $g = |g_{rs}|$. Entonces, si Δ_{rs} es el cofactor de g_{rs} en el determinante, sea $g^{pq} = \Delta_{pq}/g$. Como Riemann, introduce independientemente los cuatro símbolos índice (sin la coma)

$$(gkhi) = \frac{\partial}{\partial x_i}[gh,k] - \frac{\partial}{\partial x_h}[gi,k] + \sum_p (\{gi,p\}[hk,h] - \{gh,p\}[ik,p]).$$

Más adelante, deduce $n(n + 1)/2$ ecuaciones en derivadas parciales para las x_i como funciones de las y_i. Una típica es

$$\sum_{r,s} g_{rs} \frac{\partial x_r}{\partial y_\alpha} \frac{\partial x_s}{\partial y_\beta} = g'_{\alpha\beta}.$$

Estas ecuaciones son las condiciones necesarias y suficientes para que exista una transformación mediante la cual $F = F'$.

En parte para tratar la integrabilidad de este conjunto de ecuaciones, y en parte porque Christoffel desea considerar formas de grado superior a dos en las dx_i, lleva a cabo un número de diferenciaciones y pasos algebraicos que muestran que

$$(\alpha\delta\beta\gamma)' = \sum_{g,h,i,k} (gkhi) \frac{\partial x_g}{\partial y_\alpha} \frac{\partial x_h}{\partial y_\beta} \frac{\partial x_i}{\partial y_\gamma} \frac{\partial x_k}{\partial y_\delta}, \qquad (27)$$

donde las α, β, γ y δ toman todos los valores de 1 a n. Hay $n^2(n^2 - 1)/12$ ecuaciones de esta forma. Estas ecuaciones son las condiciones necesarias y suficientes para la equivalencia de dos formas diferenciales de cuarto orden. En efecto, sean $d^{(1)}x$, $d^{(2)}x$, $d^{(3)}x$, $d^{(4)}x$ cuatro conjuntos de diferenciales de x y de la misma manera para las y. Entonces, si tenemos la forma cuadrilineal

$$G_4 = \sum_{g,k,h,i} (gkhi) d^{(1)}x_g d^{(2)}x_k d^{(3)}x_h d^{(4)}x_i$$

las relaciones (27) son necesarias y suficientes para que $G_4 = G'_4$ donde G'_4 es el análogo de G_4 en las variables y.

Esta teoría puede ser generalizada para formas diferenciales μ-múltiples. De hecho, Christoffel introduce

$$G_\mu = \sum_{i_1,\ldots,i_\mu} (i_1 i_2 \ldots i_\mu) \underset{1}{\partial} x_{i_1} \underset{2}{\partial} x_{i_2} \ldots \underset{\mu}{\partial} x_{i_\mu} \tag{28}$$

donde el término entre paréntesis está definido en términos de las g_{rs}, así como los cuatro símbolos índice y el símbolo δ_i es utilizado para distinguir las diferenciales del conjunto de las x_i del conjunto obtenido al aplicar $\underset{j}{\partial}$. Más adelante demuestra que

$$(\alpha_1 \alpha_2 \ldots \alpha_\mu)' = \sum (i_1 \ldots i_\mu) \frac{\partial x_{i_1}}{\partial y_{\alpha_1}} \ldots \frac{\partial x_{i_\mu}}{\partial y_{\alpha_\mu}} \tag{29}$$

y obtiene condiciones necesarias y suficientes para que las G_μ sean transformables en las G'_μ.

A continuación proporciona un procedimiento general mediante el cual se puede derivar de una forma μ-ésima G una forma $(\mu + 1)$-ésima. El paso clave es introducir

$$\begin{aligned}(i_1 i_2 \ldots i_\mu i) &= \frac{\partial}{\partial x_i} (i_1 i_2 \ldots i_\mu) \\ &\quad - \sum_\lambda (\{ii_1, \lambda\}(\lambda i_2 \ldots i_\mu) + \{ii_2, \lambda\}(i_1 \lambda i_3 \ldots i_\mu) + \ldots].\end{aligned} \tag{30}$$

Estos símbolos son los coeficientes de la forma $G_{\mu+1}$. El procedimiento utilizado por Christoffel es el que Ricci y Levi-Civita llamaron más tarde diferenciación covariante (cap. 48).

Mientras que Christoffel sólo escribió un artículo clave sobre geometría riemanniana, Rudolph Lipschitz, profesor de matemáticas en la universidad de Bonn, escribió un gran número, apareciendo en el *Journal für Mathematik* a partir de 1869. Aunque hay algunas generalizaciones del trabajo de Beltrami y Christoffel, la materia esencial y los resultados son los mismos que los de los dos últimos. Obtuvo algunos nuevos resultados sobre subespacios riemannianos y espacios euclídeos n-dimensionales.

Las ideas proyectadas por Riemann y las desarrolladas por estos tres sucesores inmediatos sugirieron una gran variedad de nuevos problemas, tanto en la geometría diferencial euclídea como en la riemanniana. En particular, se generalizaron los resultados ya obte-

nidos en el caso euclídeo para tres dimensiones a curvas, superficies y formas de dimensión superior en n dimensiones. Citaremos sólo un resultado.

En 1886, Friedrich Schur (1856-1932) demostró el teorema que lleva su nombre.[16] De acuerdo con el enfoque de Riemann de la noción de curvatura, Schur habla de la curvatura de una orientación en el espacio. Tal orientación está determinada por un haz de geodésicas $\mu\alpha + \lambda\beta$ donde α y β son las direcciones de las geodésicas que pasan por un punto. Este haz forma una superficie y tiene una curvatura de Gauss que Schur llama la curvatura riemanniana de la orientación. Su teorema establece entonces que si en cada punto la curvatura riemanniana del espacio es independiente de la orientación, entonces la curvatura riemanniana es constante en todo el espacio. La variedad es entonces un espacio de curvatura constante.

5. Los invariantes de las formas diferenciales

Está claro a partir del estudio de la cuestión de cuándo una expresión dada para ds^2 puede ser transformada mediante una transformación de la forma

$$x_i = x_i(x_1', x_2', ..., x_n'), \qquad i = 1, 2, ..., n \tag{31}$$

en otra expresión semejante, con la conservación del valor de ds^2, que se obtienen diferentes representaciones coordenadas para la misma variedad. Sin embargo, las propiedades geométricas de la variedad deben ser independientes del sistema coordenado particular usando para representarla y estudiarla. Analíticamente, estas propiedades geométricas serían representadas por medio de los invariantes, esto es, expresiones que retienen su forma bajo el cambio de coordenadas y que consecuentemente tendrán el mismo valor en un punto dado. Las invariantes de interés para la geometría riemanniana conciernen no sólo la forma cuadrática fundamental, que contiene las diferenciales dx_i y dx_j, sino que también puede contener las derivadas de los coeficientes y de otras funciones. Por tanto, son llamados invariantes diferenciales.

Tomamos el caso bidimensional como ejemplo: si

[16] *Math. Ann.*, 27, 1886, 167-172 y 537-567.

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2 \tag{32}$$

es el elemento de la distancia de una superficie, entonces la curvatura gaussiana K está dada por la fórmula (8) anterior. Si ahora las coordenadas se cambian a

$$u' = f(u, v), \qquad v' = g(u, v) \tag{33}$$

entonces se tiene el teorema de que si $E\,du^2 + 2F\,du\,dv + G\,dv^2$ se transforma en $E'\,du'^2 + 2F'\,du'\,dv' + G'dv'^2$, entonces $K = K'$, donde K' es la misma expresión que en (8) pero en las variables acentuadas. De aquí que la curvatura gaussiana de una superficie sea un invariante escalar. El invariante K es llamado un invariante asociado a la forma (32) y concierne sólo a E, F y G y sus derivadas.

Lamé, en un contexto más limitado, inició el estudio de los invariantes diferenciales. Estaba interesado en invariantes bajo transformaciones de un sistema coordenado curvilíneo ortogonal en tres dimensiones a otro. Para las coordenadas cartesianas rectangulares demostró [17] que

$$\Delta_1\phi = \left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 \tag{34}$$

$$\Delta_2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} \tag{35}$$

son invariantes diferenciales (los llamó parámetros diferenciales). Así, si ϕ es transformada en $\phi'(x',y',z')$ por una transformación ortogonal (rotación de ejes) entonces

$$\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 = \left(\frac{\partial\phi'}{\partial x'}\right)^2 + \left(\frac{\partial\phi'}{\partial y'}\right)^2 + \left(\frac{\partial\phi'}{\partial z'}\right)^2$$

en el mismo punto cuyas coordenadas son (x,y,z) en el sistema original y (x',y',z') en el nuevo sistema coordenado. Se tiene la ecuación análoga para $\Delta_2\phi$.

Para las coordenadas curvilíneas ortogonales en el espacio euclídeo donde ds^2 tiene la forma

---

[17] *Jour. de l'Ecole Poly.*, 14, 1834, 191-288.

$$ds^2 = g_{11}\, du_1^2 + g_{22}\, du_2^2 + g_{33}\, du_3^2 \tag{36}$$

Lamé demostró (*Leçons sur les coordonnées curvilignes*, 1859, véase cap. 28, sec. 5) que la divergencia del gradiente de ϕ, que en coordenadas rectangulares está dado mediante el $\Delta_2\phi$ de arriba, tiene la forma invariante

$$\Delta_2\phi = \frac{1}{\sqrt{g_{11}g_{22}g_{33}}}\left[\frac{\partial}{\partial u_1}\left(\sqrt{\frac{g_{22}g_{33}}{g_{11}}}\frac{\partial\phi}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\sqrt{\frac{g_{33}g_{11}}{g_{22}}}\frac{\partial\phi}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\sqrt{\frac{g_{11}g_{22}}{g_{33}}}\frac{\partial\phi}{\partial u_3}\right)\right].$$

Incidentalmente, en este mismo trabajo proporcionó Lamé condiciones sobre cuándo las ds^2 dadas por (36) determinan un sistema coordenado curvilíneo en el espacio euclídeo y, si lo hace, cómo cambiar las coordenadas rectangulares.

La investigación de los invariantes para la teoría de superficies fue hecha primero por Beltrami,[18] que dio los dos invariantes diferenciales

$$\Delta_1\phi = \frac{1}{EG - F^2}\left\{E\left(\frac{\partial\phi}{\partial v}\right)^2 - 2F\frac{\partial\phi}{\partial u}\frac{\partial\phi}{\partial v} + G\left(\frac{\partial\phi}{\partial v}\right)^2\right\}$$

y

$$\Delta_2\phi = \frac{1}{\sqrt{EG - F^2}}\left\{\frac{\partial}{\partial u}\left(\frac{G\phi_u - F\phi_v}{\sqrt{EG - F^2}}\right) + \frac{\partial}{\partial v}\left(\frac{-F\phi_u + E\phi_v}{\sqrt{EG - F^2}}\right)\right\}.$$

Estos tienen un significado geométrico. Por ejemplo, en el caso de $\Delta_1\phi$, si $\Delta_1\phi = 1$, las curvas $\phi(u,v) =$ const. son las trayectorias ortogonales de una familia de geodésicas sobre la superficie.

La búsqueda de invariantes diferenciales se extendió a las formas diferenciales cuadráticas de n variables. De nuevo la razón fue que estos invariantes son independientes de las elecciones particulares de coordenadas; representan propiedades intrínsecas de la propia varie-

[18] *Gior. di Mat.*, 2, 1864, 267-282 y en artículos sucesivos en vols. 2 y 3 = *Opere Mat.*, 1, 107-198.

dad. De esta manera resulta que la curvatura de Riemann es un invariante escalar.

Beltrami, usando un método proporcionado por Jacobi,[19] tuvo éxito en transportar a los espacios riemannianos n-dimensionales los invariantes de Lamé.[20] Sea g como de costumbre el determinante de las g_{ij} y sea g^{ij} el cofactor dividido por g de g_{ij} en g. Entonces Beltrami demostró que el primer invariante de Lamé se convierte en

$$\Delta_1(\phi) = \sum_{i,j} g^{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_j}.$$

Esta es la forma general para el cuadrado del gradiente de ϕ. Para el segundo de los invariantes de Lamé, Beltrami obtuvo

$$\Delta_2(\phi) = \frac{1}{\sqrt{g}} \sum_i \frac{\partial}{\partial x_i} \left(\sqrt{g} \sum_i g^{ij} \frac{\partial \phi}{\partial x_j} \right).$$

También introdujo el invariante diferencial mixto

$$\Delta_1(\phi\psi) = \sum_{i,j} g^{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j}.$$

Esta es la forma general del producto escalar de los gradientes de ϕ y Ψ.

Por supuesto, la forma ds^2 es ella misma un invariante para el cambio de coordenadas. De esto, como vimos en la sección previa, Christoffel derivó formas diferenciales de orden superior, sus G_4 y G_μ, que también son invariantes. Más aún, demostró cómo de las G_μ se derivan las $G_{\mu+1}$, que también son un invariante. Lipschitz siguió también la construcción de dichos invariantes. El número y la variedad son grandes. Como veremos, esta teoría de invariantes diferenciales fue la inspiración del análisis tensorial.

---

[19] *Jour. für Math.*, 36, 1848, 113-134 = *Werke*, 2, 193-216.

[20] *Mémoire dell'Accademia delle Scienze dell'Istituto di Bologna*, (2), 8, 1868, 551-590 = *Opere Mat.*, 2, 74-118.

Bibliografía

Beltrami, Eugenio: *Opere matematiche,* 4 vols., Ulrico Hoepli, 1902-1920.

Clifford, William K.: *Mathematical papers,* Macmillan, 1882; Chelsea (reimpresión), 1968.

Coolidge, Julian L.: *A history of geometrical methods,* Dover (reimpresión), 1963, 355-387.

Encyklopadie der Mathematischen Wissenschaften, III, Teil 3, varios artículos, B. G. Teubner, 1902-1907.

Gauss, Karl F.: *Werke,* 4, 192-216, 217-258, Königliche Gesellschaft der Wissenschaften zu Göttingen, 1880. «Investigaciones generales sobre superficies curvas», la traducción al inglés ha sido reimpresa por Raven Press, 1965.

Helmholtz, Hermann von: «Über die tatsächlichen Grundlagen der Geometrie». *Wissenschaftliche Abhandlungen,* 2, 610-617.

— «Über die Tatssachen, die der Geometrie zum Grunde liegen», *Nachrichten König. Ges. der Wiss. zu Gött.,* 15, 1868, 193-221; *Wiss. Abh.,* 618-639.

— «Über den Ursprung Sinn und Bedeutung der geometrischen Sätze». Traducción al inglés: «Sobre el origen y significación de los axiomas geométricos», contenido en: Helmholtz: *Popular Scientific Papers,* Dover (reimpresión), 1962, 223-249. También en James R. Newman: *The world of mathematics,* Simon and Schuster, 1956, vol. 1, 647-668. Hay versión castellana, *Sigma. El mundo de las matemáticas,* Barcelona, Grijalbo, 8.ª ed., 1983.

Jammer, Max: *Concepts of space,* Harvard University Press, 1954.

Killing, W.: *Die nicht-euklidischen Raumformen in analytischer Behandlund,* B. G. Teubner, 1885.

Klein, F.: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert,* Chelsea (reimpresión), 1950, vol. 1, 6-62; vol. 2, 147-206.

Pierpont, James: «Some modern views of space». *Amer. Math. Soc. Bull.,* 32, 1926, 225-258.

Riemann, Bernhard: *Gesammelte mathematische Werke,* 2.ª ed. Dover (reimpresión), 1953, 272-287 y 391-404.

Russell, Bertrand: *An essay on the foundations of geometry* (1897), Dover (reimpresión), 1956.

Smith, David E.: *A source book in mathematics.* Dover (reimpresión), 1959, vol. 2, 411-425, 463-475. Contiene traducciones al inglés del ensayo de 1854 de Riemann y del ensayo de Gauss de 1822.

Staeckel, P.: «Gauss als Geometer». *Nachrichten König. Ges. der Wiss. zu Gött.,* 1917, Beiheft, 25-140; también en *Werke,* 10_2.

Weatherburn, C. E.: «The development of multidimensional differential geometry». *Australian and New Zealand Ass'n for the Advancement of Science,* 21, 1933, 12-28.

Capítulo 38
LAS GEOMETRIAS PROYECTIVA Y METRICA

> Pero siempre debiera pedirse que un asunto matemático no se considere agotado hasta que haya llegado a ser intuitivamente evidente...
>
> FELIX KLEIN

1. Introducción

Antes y durante el trabajo sobre geometría no euclídea, el estudio de las propiedades proyectivas fue la actividad geométrica principal. Además, fue evidente a partir del trabajo de von Staudt (cap. 35, sec. 3) que la geometría proyectiva precede lógicamente a la geometría euclídea, porque trata de las propiedades cualitativas y descriptivas que intervienen en la formación misma de las figuras geométricas y no usa las medidas de los segmentos de recta ni los ángulos. Este hecho sugirió que la geometría euclídea podría ser algún caso particular de la geometría proyectiva. Teniendo después a la mano las geometrías no euclídeas surgió la posibilidad de que éstas también, al menos las que tratan de los espacios de curvatura constante, pudieran ser casos especiales de la geometría proyectiva. Así, la relación de la geometría proyectiva con las geometrías no euclídeas, las cuales también son geometrías métricas, ya que emplean la distancia como un concepto fundamental, se convirtió en un tema de investigación. El aclarar la relación de la geometría proyectiva con las geometrías euclídea y no euclídeas es el gran logro de las obras que estamos a punto de examinar. Igualmente vital fue el establecimiento de la consistencia de las geometrías no euclídeas básicas.

2. Las superficies como modelos de la geometría no euclídea

Las geometrías no euclídeas que parecieron ser más significativas después de la obra de Riemann fueron las de los espacios de curvatura constante. El propio Riemann había sugerido en su artículo de 1854 que un espacio de curvatura constante positiva en dos dimensiones podría ser realizado sobre la superficie de una esfera siempre y cuando se tomara como «línea recta» la geodésica sobre la esfera. Ahora se hace referencia a esta geometría no euclídea como geometría elíptica doble por razones que se aclararán más tarde. Antes del trabajo de Riemann, la geometría no euclídea de Gauss, Lobatchevsky y Bolyai, a la cual Klein llamó después geometría hiperbólica, había sido introducida como la geometría en un plano en el que las líneas rectas ordinarias (y necesariamente infinitas) son las geodésicas. La relación de esta geometría con las variedades de Riemann no estaba clara. Riemann y Minding[1] habían pensado en las superficies de curvatura constante negativa pero ninguno de ellos relacionó éstas con la geometría hiperbólica.

Independientemente de Riemann, Beltrami reconoció[2] que las superficies de curvatura constante son espacios no euclídeos. También dio una representación limitada de la geometría hiperbólica sobre una superficie,[3] con la cual mostró que la geometría de una porción restringida del plano hiperbólico se cumple en una superficie de curvatura constante negativa si se toman como líneas rectas las geodésicas sobre esta superficie. Las longitudes y los ángulos sobre la superficie son las longitudes y los ángulos de la geometría euclídea ordinaria sobre la superficie. Una de estas superficies es conocida como la pseudoesfera (fig. 38.1). Se genera al hacer rotar una curva llamada tractriz alrededor de su asíntota. La ecuación de la tractriz es

$$z = k \log \frac{k + \sqrt{k^2 - x^2}}{x} - \sqrt{k^2 - x^2}$$

[1] *Jour. für Math.*, 19, 1839, 370-387.
[2] *Annali di Mat.*, 7, 1866, 185-204 = *Opere Mat.*, 1, 262-280.
[3] *Gior. di Mat.*, 6, 1868, 248-312 = *Opere Mat.*, 1, 374-405.

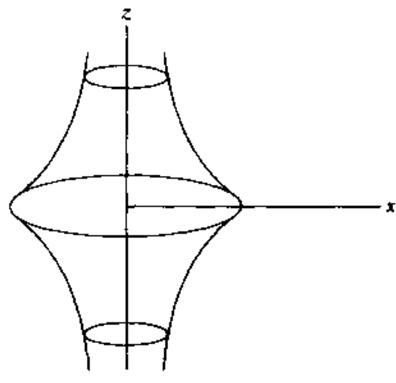


FIGURA 38.1

y la ecuación de la superficie es

$$z = k \log \frac{\sqrt{k^2 - x^2 - y^2}}{\sqrt{x^2 + y^2}} - \sqrt{k^2 - x^2 - y^2}.$$

La curvatura de la superficie es $-1/k^2$. Por tanto, la pseudoesfera es un modelo para una porción limitada del plano de Gauss, Lobatchevsky y Bolyai. Sobre la pseudoesfera se puede mover una figura y con sólo curvarla se logrará que se ajuste a la superficie, al igual que una figura plana se puede ajustar a la superficie de un cilindro circular curvándola.

Beltrami había mostrado que sobre una superficie de curvatura constante negativa se puede construir un *pedazo* del plano de Lobatchevsky. Sin embargo, no hay ninguna superficie analítica regular de curvatura constante negativa sobre la cual sea válida la geometría de *todo* el plano de Lobatchevsky. Todas esas superficies tienen una curva singular —el plano tangente no es continuo a través de ella— de modo que una continuación de la superficie a través de esta curva no continuará las figuras que representan las de la geometría de Lobatchevsky. Este resultado se debe a Hilbert.[4]

[4] *Amer. Math. Soc., Trans.*, 2, 1901, 86-99 = *Ges. Abh.*, 2, 437-448. La prueba y detalles históricos adicionales pueden encontrarse en el Apéndice V del *Grundlagen*

En relación con lo dicho, vale la pena observar que Heinrich Liebmann (1874-1939) [5] demostró que la esfera es la única superficie analítica cerrada (sin singularidades) de curvatura constante positiva, y por tanto es la única que puede utilizarse como un modelo para la geometría elíptica doble.

El desarrollo de estos modelos ayudó a los matemáticos a entender y ver el significado de las geometrías no euclídeas básicas. Se debe tener presente que estas geometrías, en el caso de dos dimensiones, son fundamentalmente geometrías del plano en las que las líneas y los ángulos son las líneas y los ángulos usuales de la geometría euclídea. Mientras la geometría hiperbólica había sido desarrollada en este estilo, las conclusiones aún parecían extrañas a los matemáticos y habían sido aceptadas sólo de mala gana dentro de las matemáticas. La geometría elíptica doble, sugerida por el enfoque de la geometría diferencial de Riemann, ni siquiera tuvo un desarrollo axiomático como una geometría del plano. Así, el único significado que los matemáticos podían ver para ella fue el proporcionado por la geometría sobre la esfera. Una comprensión mucho mejor de la naturaleza de estas geometrías fue asegurada por medio de otros desarrollos que buscaban relacionar la geometría proyectiva con la euclídea.

3. Las geometrías proyectiva y métrica

Aunque Poncelet había introducido la distinción entre propiedades proyectivas y métricas de las figuras y había establecido en su *Traité* de 1822 que las propiedades proyectivas eran lógicamente más fundamentales, fue von Staudt quien empezó a construir la geometría proyectiva sobre una base independiente de la longitud y la medida de ángulos (cap. 35, sec. 3). En 1853 Edmond Laguerre (1834-1886), profesor del Collège de France, aunque interesado principalmente en lo que sucede con los ángulos bajo una transformación proyectiva, de hecho amplió su meta al establecimiento de las

---

der Geometrie de David Hilbert, 7.ª ed., B. G. Teubner, 1930. El teorema presupone que las rectas de la geometría hiperbólica serían las geodésicas de la superficie y las longitudes y los ángulos serían las longitudes y los ángulos euclídeos en la superficie.

[5] *Nachrichten König. Ges. der wiss. zu Gött.*, 1899, 44-55; *Math. Ann.*, 53, 1900, 81-112; y 54, 1901, 505-517.

propiedades métricas de la geometría euclídea sobre la base de conceptos proyectivos, proporcionando una base proyectiva para la medida de un ángulo.[6]

Para obtener una medida del ángulo entre dos rectas dadas que se cortan, se pueden considerar dos líneas que pasan por el origen y paralelas respectivamente a las dos rectas dadas. Sean $y = x \operatorname{tg} \theta$ e $y = x \operatorname{tg} \theta'$ las ecuaciones de las dos rectas que pasan por el origen (en coordenadas no homogéneas). Sean $y = ix$ e $y = -ix$ las dos rectas (imaginarias) desde el origen a los puntos circulares en el infinito, es decir, los puntos $(1, i, 0)$ y $(1, -i, 0)$. Llamemos a estas cuatro rectas u, u', w y w' respectivamente. Sea ϕ el ángulo entre u y u'. Entonces el resultado de Laquerre es que

$$\phi = \theta' - \theta = \frac{i}{2} \log (uu', ww'), \tag{1}$$

donde (uu', ww') es la razón doble de las cuatro rectas. Lo que es significativo de la expresión (1) es que puede tomarse como la definición de la medida de un ángulo en términos del concepto proyectivo de razón doble. La función logaritmo, desde luego, es puramente cuantitativa y puede introducirse en cualquier geometría.

Independientemente de Laguerre, Cayley dio el paso siguiente. Estudió la geometría desde el punto de vista del álgebra, y de hecho estuvo interesado en la interpretación geométrica de las cuánticas (formas polinomiales homogéneas), un tema que consideraremos en el capítulo 39. Buscando mostrar que las nociones métricas pueden formularse en términos proyectivos se concentró en la relación de la geometría euclídea con la proyectiva. La obra que estamos a punto de describir se encuentra en su «Sixth Memoir upon Quantics» (Sexta memoria sobre cuánticas).[8]

La obra de Cayley resultó ser una generalización de la idea de Laguerre. Este último había usado los puntos circulares en el infinito

---

[6] *Nouvelles Annales de Mathématiques*, 12, 1853, 56-66 = *Œuvres*, 2, 6-15.

[7] La razón doble es ella misma un número complejo. El coeficiente $i/2$ asegura que un ángulo recto tenga como medida $\pi/2$. La forma en que se calculan tales razones dobles pueden encontrarse en textos sobre geometría proyectiva. Véase, por ejemplo, William C. Graustein: *Introduction to Higher Geometry*, MacMillan, 1933, cap. 8.

[8] *Phil. Trans.*, 149, 1859, 61-91 = *Coll. Math. Papers*, 2, 561-606.

para definir el ángulo en el plano. Los puntos circulares son realmente una cónica degenerada. En dos dimensiones Cayley introdujo cualquier cónica en lugar de los puntos circulares y en tres dimensiones introdujo cualquier superficie cuádrica. A estas figuras las llamó los absolutos. Cayley afirmó que todas las propiedades métricas de las figuras no son otra cosa sino las propiedades proyectivas aumentadas por el absoluto o en relación con el absoluto. Mostró entonces cómo este principio conducía a una nueva expresión para el ángulo y a una expresión para la distancia entre dos puntos.

Comienza con el hecho de que los puntos de un plano se representan mediante coordenadas homogéneas. Estas coordenadas no tienen que ser vistas como distancias o razones de distancias sino como una noción fundamental aceptada que no requiere ni admite explicación. Para definir la distancia y la medida de un ángulo introduce la forma cuadrática

$$F(x, x) = \sum_{i,j=1}^{3} a_{ij}x_ix_j, \qquad a_{ij} = a_{ji}$$

y la forma bilineal

$$F(x, y) = \sum_{i,j=1}^{3} a_{ij}x_iy_j.$$

La ecuación $F(x,x) = 0$ define una cónica que es el absoluto de Cayley. La ecuación del absoluto en coordenadas de rectas es

$$G(u, u) = \sum_{i,j=1}^{3} A^{ij}u_iu_j = 0,$$

donde A^{ij} *es el cofactor de* a_{ij} *en el determinante* $|a|$ de los coeficientes de F.

Cayley define entonces la distancia δ entre dos puntos x y y, donde $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$, por medio de la fórmula

$$\delta = \text{arc cos} \frac{F(x, y)}{[F(x, x)F(y, y)]^{1/2}}. \tag{2}$$

El ángulo ϕ entre dos rectas cuyas coordenadas de rectas son $u = (u_1, u_2, u_3)$ y $v = (v_1, v_2, v_3)$ se define por

$$\cos\phi = \frac{G(u, v)}{[G(u, u)G(v, v)]^{1/2}}. \tag{3}$$

Estas fórmulas generales se hacen más sencillas si tomamos como absoluto la cónica particular $x_1{}^2 + x_2{}^2 + x_3{}^2 = 0$. Entonces, si (a_1, a_2, a_3) y (b_1, b_2, b_3) son las coordenadas homogéneas de dos puntos, la distancia entre ellos está dada por

$$\text{arc cos}\, \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}} \tag{4}$$

y el ángulo ϕ entre dos rectas cuyas coordenadas homogéneas de rectas son (u_1, u_2, u_3) y (v_1, v_2, v_3) está dado por

$$\cos\phi = \frac{u_1v_1 + u_2v_2 + u_3v_3}{\sqrt{u_1^2 + u_2^2 + u_3^2}\sqrt{v_1^2 + v_2^2 + v_3^2}}. \tag{5}$$

Con respecto a la expresión para la distancia, si usamos la simplificación de que $xy = x_1y_1 + x_2y_2 + x_3y_3$ entonces, si $a = (a_1, a_2, a_3)$, b y c son tres puntos sobre una recta,

$$\text{arc cos}\, \frac{ab}{\sqrt{aa}\sqrt{bb}} + \text{arc cos}\, \frac{bc}{\sqrt{bb}\sqrt{cc}} = \text{arc cos}\, \frac{ac}{\sqrt{aa}\sqrt{cc}}.$$

Esto es, las distancias se suman como se debe. Tomando como cónica absoluto los puntos circulares en el infinito $(1, i, 0)$ y $(1, -i, 0)$, Cayley demostró que sus fórmulas para la distancia y el ángulo se reducen a las fórmulas euclídeas usuales.

Se observará que las expresiones para la longitud y el ángulo involucran la expresión algebraica para el absoluto. Generalmente la expresión analítica para cualquier propiedad métrica euclídea involucra la relación de esa propiedad con el absoluto. Las propiedades métricas no son propiedades de la figura *per se*, sino de la figura en relación con el absoluto. Esta es la idea de Cayley de la determinación proyectiva general de la métrica. El lugar del concepto métrico en la geometría proyectiva y la generalidad mayor de la última fueron descritos por Cayley como «La geometría métrica es parte de la geometría proyectiva».

La idea se Cayley fue retomada por Felix Klein (1849-1925) y generalizada de modo que incluyera las geometrías no euclídeas.

Klein, profesor de Göttingen, fue una de las figuras de las matemáticas en Alemania durante la última parte del siglo XIX y la primera parte del XX. Durante los años 1869-1870 estudió el trabajo de Lobatchevsky, Bolyai, von Staudt y Cayley; sin embargo, aún en 1871 no conocía el resultado de Laguerre. A él le parecía que era posible subsumir las geometrías no euclídeas, la geometría elíptica doble y la hiperbólica, en la geometría proyectiva explotando la idea de Cayley. Proporcionó un bosquejo de sus ideas en un artículo de 1871 [9] y después los desarrolló en otros dos artículos.[10] Klein fue el primero en reconocer que no necesitamos superficies para obtener modelos de las geometrías no euclídeas.

Para empezar, Klein observó que Cayley no dejó claro exactamente lo que consideraba como significado de sus coordenadas. Eran simplemente variables sin interpretación geométrica alguna o eran distancias euclídeas. Pero para obtener las geometrías métricas a partir de la geometría proyectiva era necesario construir las coordenadas sobre una base proyectiva. Von Staudt había demostrado (cap. 35, sec. 3) que era posible asignar números a los puntos por medio de su álgebra de lanzamientos. Pero usó el axioma euclídeo del paralelismo. A Klein le parecía claro que se podía prescindir de este axioma y en el artículo de 1873 muestra que esto puede hacerse. Por tanto, las coordenadas y la razón doble de cuatro puntos, cuatro rectas o cuatro planos puede definirse sobre una base puramente proyectiva.

La idea más importante de Klein fue que por medio de la particularización de la naturaleza de la superficie cuádrica absoluto de Cayley (si se considera la geometría tridimensional), podría mostrarse que la métrica, la cual según Cayley dependía de la naturaleza del absoluto, conduciría a la geometría hiperbólica y a la elíptica doble. Cuando la superficie de segundo grado es un elipsoide real, un paraboloide elíptico real o un hiperboloide real de dos hojas, se obtiene la geometría métrica de Lobatchevsky. Cuando la superficie de segundo grado es imaginaria, se obtiene la geometría no euclídea de Riemann (de curvatura constante positiva). Si el absoluto es la esfera-círculo, cuya ecuación en coordenadas homogéneas es

---

[9] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1871, 419-433 = *Ges. Math. Abh.*, 1, 244-253.

[10] *Math. Ann.*, 4, 1871, 573-625; y 6, 1873, 112-145 = *Ges. Math. Abh.*, 1, 254-305, 311-343.

$x^2 + y^2 + z^2 = O$, $t = O$, entonces se obtiene la geometría métrica euclídea usual. Así las geometrías métricas se convierten en casos especiales de la geometría proyectiva.

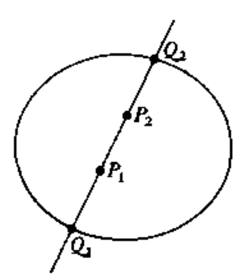


FIGURA 38.2

Para apreciar las ideas de Klein consideremos la geometría de dos dimensiones. Se escoge una cónica en el plano proyectivo; esta cónica será el absoluto. Su ecuación es

$$F = \sum_{i,j=1}^{3} a_{ij}x_ix_j = 0 \tag{6}$$

en coordenadas puntuales y

$$G = \sum_{i,j=1}^{3} A^{ij}u_iu_j = 0 \tag{7}$$

en coordenadas de rectas. Para obtener la geometría de Lobatchevsky, la cónica debe ser real, por ejemplo, en coordenadas homogéneas planas $x_1^2 + x_2^2 - x_3^2 = O$; para la geometría de Riemann sobre una superficie de curvatura constante positiva es imaginaria, por ejemplo, $x_1^2 + x_2^2 + x_3^2 = O$; y para la geometría euclídea, la cónica degenera en dos rectas que coinciden, representadas en coordenadas homogéneas por $x_3 = O$ y sobre este lugar geométrico se escogen dos puntos imaginarios cuya ecuación es $x_1^2 + x_2^2 = O$, esto es, los puntos circulares del infinito cuyas coordenadas homogéneas son $(1, i, O)$ y $(1, -i, O)$. En todos los casos la cónica tiene una ecuación real.

Para ser precisos supongamos que la cónica es la que se muestra en la figura 38.2. Si P_1 y P_2 son dos puntos de una recta, esta recta corta al absoluto en dos puntos (reales o imaginarios). Entonces la distancia se toma como

$$d = c \log (P_1P_2, Q_1Q_2), \tag{8}$$

donde la cantidad entre paréntesis denota la razón doble de los cuatro puntos y c es una constante. Esta razón doble puede expresarse en términos de las coordenadas de los puntos. Es más, si hay tres puntos P_1, P_2, P_3 en la recta entonces puede mostrarse fácilmente que

$$(P_1P_2, Q_1Q_2) \cdot (P_2P_3, Q_1Q_2) = (P_1P_3, Q_1Q_2)$$

de modo que $P_1\, P_2 + P_2P_3 = P_1P_3$.

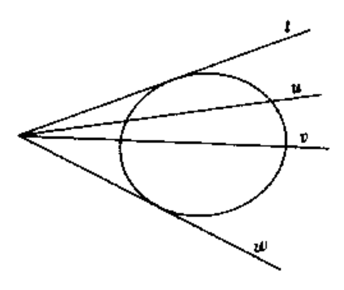


FIGURA 38.3

Análogamente, si u y v son dos rectas (fig. 38.3) se consideran las tangentes t y w desde su punto de intersección al absoluto (las tangentes pueden ser rectas imaginarias); entonces el ángulo entre u y v se define como

$$\phi = c' \log (uv, tw),$$

donde nuevamente c' es una constante y la cantidad entre paréntesis denota la razón doble de las cuatro rectas.

Para expresar los valores de d y ϕ analíticamente y demostrar su dependencia de la elección del absoluto, sea dada la ecuación del absoluto por las F y G anteriores. Por definición

$$F_{xy} = \sum_{i,j=1}^{3} a_{ij}x_iy_j.$$

Ahora puede demostrarse que si $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ son las coordenadas de P_1 y P_2 entonces

$$d = c \log \frac{F_{xy} + \sqrt{F_{xy}^2 - F_{xx}F_{yy}}}{F_{xy} - \sqrt{F_{xy}^2 - F_{xx}F_{yy}}}.$$

Análogamente, si (u_1, u_2, u_3) y (v_1, v_2, v_3) son las coordenadas de las dos rectas, entonces puede demostrarse, usando G, que

$$\phi = c' \log \frac{G_{uv} + \sqrt{G_{uv}^2 - G_{uu}G_{vv}}}{G_{uv} - \sqrt{G_{uv}^2 - G_{uu}G_{vv}}}.$$

La constante c' se toma generalmente como $i/2$, de modo que ϕ sea real y un ángulo central completo mida 2π.

Klein usó las expresiones logarítmicas anteriores para el ángulo y la distancia y demostró cómo pueden obtenerse las geometrías métricas a partir de la geometría proyectiva. Es decir, que si se parte de la geometría proyectiva, entonces por la elección del absoluto y usando las expresiones anteriores para la distancia y el ángulo pueden obtenerse las geometrías euclídea, hiperbólica y elíptica como casos especiales. La naturaleza de la geometría métrica queda determinada por la elección del absoluto. Incidentalmente, puede mostrarse que las expresiones de Klein para la distancia y el ángulo son iguales a las de Cayley.

Si se realiza una transformación proyectiva (*i.e.*, lineal) del plano proyectivo en sí mismo, la cual transforma el absoluto en sí mismo (aunque los puntos en el absoluto vayan a otros puntos), entonces, debido a que la razón doble queda inalterada por una transformación lineal, las distancias y los ángulos tampoco se alteran. Estas transformaciones lineales particulares que dejan fijo el absoluto son los movimientos rígidos o transformaciones de congruencia de la geometría métrica particular determinada por el absoluto. Una transformación proyectiva general no dejará invariante el absoluto. Así la propia geometría proyectiva es realmente más general por las transformaciones que permite.

Otra contribución de Klein a la geometría no euclídea fue la observación, que él dice hizo primero en 1871 [11], pero que publicó en 1874,[12] de que hay dos clases de geometría elíptica. En la geo-

[11] *Math. Ann.*, 4, 1871, 604. Véase también *Math. Ann.*, 6, 1873, 125; y *Math. Ann.*, 37, 1890, 554-557.

[12] *Math. Ann.*, 7, 1874, 549-557; 9, 1876, 476-482 = *Ges. Math. Abh.*, 2, 63-77.

metría elíptica doble dos puntos no siempre determinan una línea recta única. Esto es evidente a partir del modelo esférico cuando los dos puntos son diametralmente opuestos. En la segunda geometría elíptica, llamada elíptica simple, dos puntos siempre determinan una línea recta única. Cuando la investigó desde el punto de vista de la geometría diferencial, la forma diferencial ds^2 de una superficie de curvatura constante positiva es (en coordenadas homogéneas)

$$ds^2 = \frac{dx_1^2 + dx_2^2 + dx_3^2}{\{1 + (a^2/4)(x_1^2 + x_2^2 + x_3^2)\}^2} .$$

En ambos casos $a^2 > 0$. Sin embargo, en el primer tipo las geodésicas son curvas de longitud finita $2\pi/a$ o, si R es el radio, $2\pi R$, y son cerradas (regresan sobre sí mismas). En el segundo las geodésicas son de longitud π/a o πR y también son cerradas.

Un modelo de una superficie que tiene las propiedades de la geometría elíptica simple, y que se le debe a Klein,[13] es proporcionado por un hemisferio que incluye la frontera. Sin embargo, deben identificarse dos puntos cualesquiera sobre la frontera que sean diametralmente opuestos. Los arcos circulares máximos en la semiesfera son las «líneas rectas» o geodésicas de esta geometría y los ángulos ordinarios en la superficie son los ángulos de la geometría. La geometría elíptica simple (a veces llamada elíptica, en cuyo caso la geometría elíptica doble es llamada esférica) también es realizada entonces en un espacio de curvatura constante positiva. Realmente no pueden unirse los pares de puntos que son identificados en este modelo, al menos en el espacio tridimensional. La superficie tendría que cruzarse a sí misma y los puntos que coincidieran en la intersección tendrían que ser considerados distintos.

Ahora podemos ver por qué Klein introdujo la terminología de hiperbólica para la geometría de Lobatchevsky, elíptica para el caso de la geometría de Riemann sobre una superficie de curvatura constante positiva, y parabólica para la geometría euclídea. Esta terminología fue sugerida por el hecho de que la hipérbola ordinaria interseca la línea del infinito en dos puntos y correspondientemente en la geometría hiperbólica cada línea interseca al absoluto en dos puntos reales. La elipse ordinaria no tiene puntos reales en común con la recta del infinito y en la geometría elíptica, de modo parecido,

[13] Véanse las referencias en la nota 11.

toda recta carece de puntos reales en común con el absoluto. La parábola ordinaria solamente tiene un punto real en común con la recta del infinito y en la geometría euclídea (cómo es extendida en la geometría proyectiva) cada recta tiene un punto real en común con el absoluto.

El sentido que gradualmente surgió de las contribuciones de Klein fue que la geometría proyectiva es en realidad lógicamente independiente de la geometría euclídea. Más aún, las geometrías no euclídeas y euclídea también fueron vistas como casos especiales o subgeometrías de la geometría proyectiva. De hecho, la obra estrictamente lógica o rigurosa sobre los fundamentos axiomáticos de la geometría proyectiva y sus relaciones con las subgeometrías quedaba por realizarse (cap. 42). Pero haciendo ver claro el papel básico de la geometría proyectiva, Klein abrió el camino a un desarrollo axiomático que podría comenzar con la geometría proyectiva y obtener de ella las varias geometrías métricas.

4. Los modelos y el problema de la consistencia

Al principio de los 1870 se habían introducido y estudiado intensivamente varias geometrías no euclídeas básicas, la hiperbólica y las dos geometrías elípticas. La cuestión fundamental que aún quedaba por resolver para lograr que estas geometrías fueran ramas legítimas de las matemáticas consistía en determinar si eran consistentes. Toda la obra realizada por Gauss, Lobatchevsky, Bolyai, Riemann, Cayley y Klein podría haber resultado un sinsentido si hubiera contradicciones inherentes a estas geometrías.

De hecho, la prueba de la consistencia de la geometría elíptica doble en dos dimensiones estaba a mano, y posiblemente Riemann apreció este hecho, aunque no hizo afirmación explícita alguna. Beltrami [14] había señalado que la geometría de Riemann en dos dimensiones con curvatura constante positiva se construye sobre una esfera. Este modelo hace posible la demostración de la consistencia de la geometría elíptica doble en dos dimensiones. Los axiomas (que no eran explícitos en esa época) y los teoremas de esta geometría son todos aplicables a la geometría de la superficie de la esfera siempre que se interprete la recta en la geometría elíptica doble como el

[14] *Annali di Mat.* (2), 2, 1868-1869, 232-255 = *Opere Matematiche*, 1, 406-429.

círculo máximo sobre la superficie de la esfera. Si hubiera teoremas contradictorios en esta geometría elíptica doble entonces habría teoremas contradictorios en la geometría de la superficie de la esfera. Ahora bien, la esfera es parte de la geometría euclídea. Por tanto, si la geometría euclídea es consistente, entonces la geometría elíptica doble también debe serlo. Para los matemáticos de los años 1870 la consistencia de la geometría euclídea difícilmente estaba sujeta a dudas porque, aparte de los puntos de vista de unos pocos autores como Gauss, Bolyai, Lobatchevsky y Riemann, la geometría euclídea era todavía la geometría necesaria del mundo físico y era inconcebible que pudiera haber propiedades contradictorias en la geometría del mundo físico. Sin embargo, es importante, especialmente a la luz de desarrollos posteriores, darse cuenta de que esta demostración de la consistencia de la geometría elíptica doble depende de la consistencia de la geometría euclídea.

El método para demostrar la consistencia de la geometría elíptica doble no podía utilizarse para la geometría elíptica simple o para la geometría hiperbólica. El modelo hemisférico para la geometría elíptica simple no puede ser llevado a cabo en la geometría euclídea tridimensional aunque sí en la geometría euclídea de cuatro dimensiones. Si uno estuviera dispuesto a creer en la consistencia de esta última, entonces podría aceptar la consistencia de la geometría elíptica simple. Sin embargo, aunque la geometría n-dimensional ya había sido considerada por Grassmann, Riemann y otros, es dudoso que cualquier matemático de los años 1870 hubiese estado dispuesto a afirmar la consistencia de la geometría euclídea de cuatro dimensiones.

El caso de la consistencia de la geometría hiperbólica no podría llevarse a cabo sobre alguna de estas bases. Beltrami había dado la interpretación de la pseudoesfera, que es una superficie en el espacio euclídeo, pero esto sirve como modelo sólo para una región limitada de la geometría hiperbólica y, por tanto, no pudo usarse para establecer la consistencia de toda la geometría. Lobatchevsky y Bolyai habían considerado este problema (cap. 36, sec. 8), pero no pudieron aclararlo. El caso es que aunque Bolyai publicó orgullosamente su geometría no euclídea, hay pruebas de que dudó de su consistencia, ya que en artículos que se encontraron después de su muerte, continuó tratando de demostrar el axioma euclídeo del paralelismo.

La consistencia de las geometrías hiperbólica y elíptica simple fue establecida por medio de nuevos modelos. El modelo para la geo-

metría hiperbólica es debido a Beltrami.[15] Sin embargo, la función distancia que se usa en este modelo es debida a Klein y a él se adscribe frecuentemente el modelo. Consideremos el caso bidimensional.

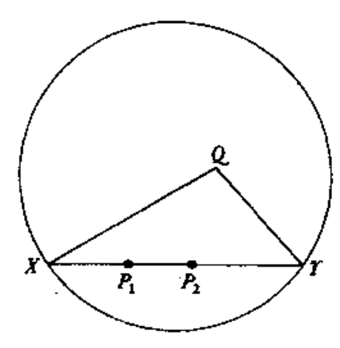


FIGURA 38.4

Dentro del plano euclídeo (que es parte del plano proyectivo) se elige una cónica real que, indistintamente, puede ser el círculo (fig. 38.4). De acuerdo con esta representación de la geometría hiperbólica, los puntos de la geometría son los puntos interiores de este círculo. Una recta de esta geometría es una cuerda del círculo, digamos la cuerda XY (sin incluir a X e Y). Si tomamos cualquier punto Q que no esté en XY entonces podemos encontrar cualquier número de rectas que pasen por Q y que no corten a XY. Dos de estas rectas, digamos, QX y QY, separan las rectas que pasan por Q en dos clases, aquellas que cortan a XY y a las que no la cortan. En otras palabras, el axioma del paralelismo de la geometría hiperbólica se satisface por los puntos y rectas (cuerdas) interiores del círculo. Además, sea

$$\measuredangle (a, b) = \frac{1}{2i} \log (ab, mn),$$

la medida del ángulo formado por dos rectas a y b, donde m y n son las tangentes imaginarias conjugadas desde el vértice del ángulo

[15] *Annali di Mat.*, 7, 1886, 185-204 = *Opere Mat.*, 1, 262-280; *Gior. di Mat.*, 6, 1868, 284-312 = *Opere Mat.*, 1, 374-405.

al círculo y (ab,mn) es la razón doble de las cuatro rectas a, b, m y n. La constante $1/2i$ asegura que un ángulo recto tenga la medida $\pi/2$. La definición de la distancia entre dos puntos está dada por la fórmula (8), esto es, $d = c \log (PP',XY)$ tomándose generalmente c como $k/2$. De acuerdo con esta fórmula, conforme P o P' se acercan a X o Y, la distancia PP' se hace infinita. De este modo, en términos de esta distancia, una cuerda es una recta infinita de la geometría hiperbólica.

Así, con las definiciones proyectivas de distancia y medida de ángulos, los puntos, cuerdas, ángulos y otras figuras en el interior del círculo satisfacen los axiomas de la geometría hiperbólica. Entonces los teoremas de la geometría hiperbólica se aplican también a estas figuras dentro del círculo. En este modelo los axiomas y teoremas de la geometría hiperbólica son realmente afirmaciones acerca de figuras y conceptos especiales (*e.g.*, la distancia definida a la manera de la geometría hiperbólica) de la geometría *euclídea*. Como los axiomas y teoremas en cuestión se aplican a estas figuras y conceptos, vistos como pertenecientes a la geometría euclídea, todas las afirmaciones de la geometría hiperbólica son teoremas de la geometría euclídea. Entonces, si hubiese alguna contradicción en la geometría hiperbólica, esta contradicción sería una contradicción dentro de la geometría euclídea. Pero *si la geometría euclídea es consistente*, entonces la geometría hiperbólica también debe serlo. Así la consistencia de la geometría hiperbólica se reduce a la consistencia de la geometría euclídea.

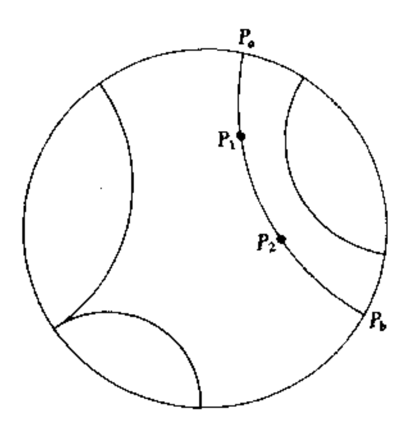


FIGURA 38.5

El hecho de que la geometría hiperbólica sea consistente implica que el axioma euclídeo del paralelismo es independiente de los otros axiomas euclídeos. Si este no fuera el caso, esto es, si el axioma euclídeo del paralelismo se pudiera obtener de los otros axiomas, sería también un teorema de la geometría hiperbólica, pues además del axioma del paralelismo, los otros axiomas de la geometría euclídea son los mismos que los de la geometría hiperbólica. Pero este teorema contradiría el axioma del paralelismo de la geometría hiperbólica y entonces ésta sería inconsistente. La consistencia de la geometría elíptica simple en dos dimensiones puede demostrarse de la misma manera que para la geometría hiperbólica, porque esta geometría elíptica también se realiza dentro del plano proyectivo y con la definición proyectiva de distancia.

Independientemente y en relación con su trabajo sobre funciones automorfas, Poincaré [16] dio otro modelo que también establece la consistencia de la geometría hiperbólica. Una forma en la cual este modelo de Poincaré para la geometría hiperbólica plana [17] puede expresarse toma el absoluto como un círculo (fig. 38.5). Dentro del absoluto las líneas rectas de la geometría son arcos de círculo que cortan al absoluto ortogonalmente y líneas rectas que pasan por el centro del absoluto. La longitud de cualquier segmento P_1P_2 está dada por log (P_1P_2,P_aP_b), donde $(P_1P_2,P_aP_b) = (P_1P_b/P_2P_b)/(P_1P_a/P_2P_a)$, P_a y P_b son los puntos en los cuales el arco que pasa por P_1 y P_2 corta al absoluto, y las longitudes P_1P_b, P_2P_b, etc. son las cuerdas. El ángulo entre dos «rectas» de este modelo que se intersecan es el ángulo euclídeo normal entre los dos arcos. Dos arcos circulares que sean tangentes en un punto en el absoluto son «rectas» paralelas. Puesto que en este modelo los axiomas y teoremas de la geometría hiperbólica también son teoremas especiales de la geometría euclídea, el argumento dado antes a propósito del modelo de Beltrami puede aplicarse aquí para establecer la consistencia de la geometría hiperbólica. Los análogos en dimensiones mayores de los modelos anteriores son también válidos.

---

[16] *Acta Math.*, 1, 1882, 1-62 = *Œuvres*, 2, 108-168; véase pág. 8 y pág. 52 del artículo.

[17] Esta forma, atribuida a Poincaré, es parecida a una que él dio en el *Bull. Soc. Math. de France*, 15, 1887, 203-216 = *Œuvre, 11, 79-91. El modelo descrito aquí parece haber sido dado primero por Joseph Wellstein (1869-1919) en H. Weber y J. Wellstein, Enzyklopädie der Elementar-Mathematik*, 2, 1905, 39-81.

5. La geometría desde el punto de vista de las transformaciones

El éxito de Klein en subsumir las varias geometrías métricas en la geometría proyectiva lo condujo a buscar el caracterizar las varias geometrías no solamente sobre la base de las propiedades no métricas y métricas y las distinciones entre las métricas, sino desde el punto de vista más amplio de lo que estas geometrías, y otras geometrías que ya habían aparecido en escena, intentaban lograr. Proporcionó esta caracterización en un discurso de 1872, «Vergleichende Betrachtungen über neuere geometrische Forschungen» (Una Revisión Comparativa de las Investigaciones Recientes en Geometría),[18] con ocasión de su admisión en la facultad de la universidad de Erlangen, y las opiniones allí expresadas se conocen como el *Erlanger Programm*.

La idea básica de Klein es que cada geometría puede ser caracterizada por un grupo de transformaciones y que una geometría trata realmente de los invariantes por este grupo de transformaciones. Más aún, una subgeometría de una geometría es la colección de los invariantes bajo un subgrupo de transformaciones del grupo original. Con esta definición todos los teoremas de una geometría correspondientes a un grupo dado continúan siendo teoremas en la geometría del subgrupo.

Aunque Klein no da en su artículo las formulaciones analíticas de los grupos de transformaciones que discute, daremos algunas para ser explícitos. De acuerdo con su noción de una geometría, la geometría proyectiva, digamos en dos dimensiones, es el estudio de los invariantes bajo el grupo de transformaciones desde los puntos de un plano a los de otro plano, o a los del mismo plano (colineaciones). Cada transformación es de la forma

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned} \tag{9}$$

en donde se presuponen las coordenadas homogéneas, los a_{ij} son números reales y el determinante de los coeficientes debe ser dife-

[18] *Math. Ann.*, 43, 1893, 63-100 = *Ges. Math. Abh.*, 1, 460-497. Una traducción al inglés puede encontrarse en el *N. Y. Math. Soc. Bull.*, 2, 1893, 215-249.

rente de cero. En coordenadas no homogéneas las transformaciones se representan por

$$x' = \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}$$
$$y' = \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}} \quad (10)$$

y nuevamente el determinante de los a_{ij} debe ser diferente de cero. Los invariantes bajo el grupo proyectivo son, por ejemplo, linealidad, colinealidad, razón doble, conjuntos armónicos y la propiedad de ser una sección cónica.

Un subgrupo del grupo proyectivo es la colección de las transformaciones afines.[19] Este subgrupo se define como sigue: fíjese una recta l_∞ cualquiera en el plano proyectivo. Los puntos de l_∞ son llamados puntos ideales o puntos del infinito y l_∞ es llamada la recta del infinito. Los otros puntos y rectas del plano proyectivo son llamados puntos ordinarios y estos son los puntos usuales del plano euclídeo. El grupo afín de colineaciones es el subgrupo del grupo proyectivo que deja invariante l_∞ (aunque no necesariamente punto a punto) y la geometría afín es el conjunto de propiedades y relaciones invariantes por el grupo afín. Algebraicamente, en dos dimensiones y en coordenadas homogéneas, las transformaciones afines se representan por las ecuaciones (9) anteriores, pero en las cuales $a_{31} = a_{32} = 0$, y con la misma condición para el determinante. En coordenadas no homogéneas las transformaciones afines están dadas por

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13} \\ y' &= a_{21}x + a_{22}y + a_{23} \end{aligned} \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Bajo una transformación afín, las líneas rectas van a líneas rectas y las líneas rectas paralelas a rectas paralelas. Sin embargo, las longitudes y medidas de ángulos se alteran. La geometría afín fue observada primero por Euler y después por Möbius en su *Der barycentrische Calcul*. Es útil en el estudio de la mecánica de las deformaciones.

[19] Klein no señaló este subgrupo.

El grupo de cualquier geometría métrica es el mismo que el grupo afín, excepto que el determinante anterior debe tener el valor +1 o −1. La primera de las geometrías métricas es la geometría euclídea. Para definir al grupo de esta geometría se empieza con l_∞ y se supone que hay una involución fija sobre l_∞. Se requiere que esta involución no tenga puntos reales dobles sino que tenga a los puntos circulares en el ∞ como puntos dobles (imaginarios). Luego consideramos todas las transformaciones proyectivas que no solamente dejan fija la l_∞, sino que llevan cualquier punto de la involución al punto correspondiente de ella, lo cual implica que cada punto circular va sobre sí mismo. Algebraicamente estas transformaciones del grupo euclídeo se representan en coordenadas no homogéneas (bidimensionales) por

$$\begin{aligned} x' &= \varrho(x \cos\theta - y \operatorname{sen}\theta + \alpha) \\ y' &= \varrho(x \operatorname{sen}\theta + y \cos\theta + \beta) \end{aligned}, \qquad \varrho = \pm 1.$$

Los invariantes son longitud, medida de ángulos y medida y forma de cualquier figura.

La geometría euclídea, según se usa el término en esta clasificación, es el conjunto de invariantes bajo esta clase de transformaciones. Las transformaciones son rotaciones, traslaciones y reflexiones. Para obtener los invariantes asociados con figuras semejantes, se introduce el subgrupo del grupo afín conocido como el grupo métrico parabólico. Este grupo se define como la clase de las transformaciones proyectivas que dejan invariantes la involución sobre l_∞, y esto significa que cada par de puntos correspondientes va a algún par de puntos correspondientes. En coordenadas no homogéneas las transformaciones del grupo métrico parabólico son de la forma

$$\begin{aligned} x' &= ax - by + c \\ y' &= bex + aey + d \end{aligned}$$

donde $a^2 + b^2 \neq 0$, y $e^2 = 1$. Estas transformaciones conservan la medida de los ángulos.

Para caracterizar la geometría métrica hiperbólica regresamos a la geometría proyectiva y consideramos una cónica no degenerada real arbitraria (el absoluto) en el plano proyectivo. El subgrupo del

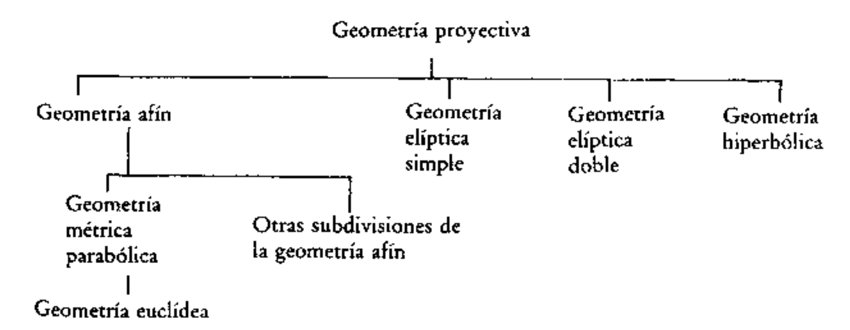
grupo proyectivo que deja esta cónica invariante (aunque no necesariamente punto a punto) es llamado grupo métrico hiperbólico y la geometría correspondiente es la geometría métrica hiperbólica. Los invariantes son aquellos que están asociados con la congruencia.

La geometría elíptica simple es la geometría correspondiente al subgrupo de las transformaciones proyectivas que deja invariante una elipse imaginaria dada (el absoluto) en el plano proyectivo. El plano de la geometría elíptica es el plano proyectivo real y los invariantes son aquellos asociados con la congruencia.

Incluso la geometría elíptica doble puede abarcarse desde este punto de vista de las transformaciones, pero se debe comenzar con las transformaciones proyectivas en tres dimensiones para caracterizar la geometría métrica bidimensional. El subgrupo de transformaciones consiste en aquellas transformaciones proyectivas tridimensionales que transforman una esfera dada (superficie) *S* de la porción finita del espacio en sí misma. La superficie esférica *S* es el «plano» de la geometría elíptica doble. Nuevamente los invariantes son aquellos asociados con la congruencia.

En cuatro geometrías métricas, esto es, la euclídea, la hiperbólica y las dos geometrías elípticas, las transformaciones que son permitidas en el correspondiente subgrupo son las que generalmente se llaman movimientos rígidos y estas son las únicas geometrías que permiten movimientos rígidos.

Klein introdujo varias clasificaciones intermedias que no repetiremos aquí. El esquema que sigue muestras las relaciones de las principales geometrías.



Klein también consideró geometrías más generales que la proyectiva. En esta época (1872) se estaba distinguiendo gradualmente

la geometría algebraica como una disciplina separada y caracterizó a esta geometría introduciendo las transformaciones que en tres dimensiones y en coordenadas no homogéneas se escriben como

$$x' = \phi(x, y, z), y' = \psi(x, y, z), z' = \chi (x, y, z).$$

Requirió que las funciones ϕ, ψ y χ fueran racionales y univaluadas y que debería ser posible resolver las ecuaciones para x, y y z en términos de funciones racionales univaluadas de x', y' y z'. Estas transformaciones son llamadas transformaciones de Cremona y los invariantes por ellas son el objeto de estudio de la geometría algebraica (cap. 39).

Klein también proyectó el estudio de los invariantes bajo transformaciones continuas uno-a-uno con inversas continuas. Esta es la clase ahora llamada de los homeomorfismos y el estudio de los invariantes bajo tales transformaciones es el objeto de estudio de la topología (cap. 50). Aunque Riemann también había considerado lo que ahora se reconoce que son problemas topológicos en su trabajo con las superficies de Riemann, la previsión de la topología como una geometría importante fue un paso audaz en 1872.

Desde los días de Klein, ha sido posible llevar a cabo adiciones y especializaciones más amplias en la clasificación de Klein. Pero no toda la geometría puede ser incorporada en el esquema de Klein. Hoy, la geometría algebraica y la geometría diferencial no caen dentro de este esquema.[20] A pesar de que el enfoque de Klein de la geometría no demostró abarcar todo, sí proporcionó un método sistemático de clasificación y estudio de una buena parte de la geometría y sugirió numerosos problemas de investigación. Su «definición» de geometría guió el pensamiento geométrico durante aproximadamente cincuenta años. Más aún, su énfasis en los invariantes por las transformaciones fue más allá de las matemáticas, generalmente a la mecánica y a la física matemática. El problema físico de la invariancia bajo transformaciones o el problema de expresar las leyes físicas de una manera independiente del sistema de coordenadas fue de mucha importancia en el pensamiento físico después de que se observó la invariancia de las ecuaciones de Maxwell bajo las transfor-

[20] Klein sí habla, en el caso de la geometría diferencial, del grupo de transformaciones que dejan invariante la expresión ds^2. Esto conduce a los invariantes diferenciales (*Ges. math. Abh.*, 1, 487).

maciones de Lorentz (un subgrupo de cuatro dimensiones de la geometría afín). Esta línea de pensamiento condujo hacia la teoría especial de la relatividad.

Mencionaremos aquí meramente estudios más avanzados sobre la clasificación de las geometrías que, al menos en su época, atrajeron considerable atención. Helmholtz y Sophus Lie (1842-1899) buscaron caracterizar las geometrías en las que son posibles los movimientos rígidos. El artículo básico de Helmholtz, «Über die Thatsachen, die der Geometrie zum Grunde liegen» (Sobre los hechos que subyacen a la Geometría),[21] mostró que si los movimientos de los cuerpos rígidos tienen que ser posibles en un espacio, entonces la expresión de Riemann para *ds* en un espacio de curvatura constante es la única posible. Lie atacó el mismo problema por medio de lo que se conoce como la teoría de los grupos de transformaciones continuas, una teoría que él ya había introducido en el estudio de las ecuaciones diferenciales ordinarias, y caracterizó los espacios en los que los movimientos rígidos son posibles por medio de los tipos de grupos de transformaciones que permiten estos espacios.[22]

6. La realidad de la geometría no euclídea

El interés en las geometrías no euclídeas sintéticas clásicas y en la geometría proyectiva declinó después de la obra de Klein y Lie, en parte debido a que la esencia de estas estructuras fue expuesta de modo muy claro desde el punto de vista de las transformaciones. El sentimiento de los matemáticos en cuanto concierne al descubrimiento de teoremas adicionales era que la mina había sido agotada. La rigorización de los fundamentos quedó por hacer y esta fue una área activa durante muchos años después de 1880 (cap. 42).

Otra razón de la pérdida de interés en las geometrías no euclídeas fue su aparente falta de relevancia para el mundo físico. Es curioso que los primeros que se dedicaron a este campo, Gauss, Lobatchevsky y Bolyai sí pensaban que la geometría no euclídea podría ser aplicable cuando se hubiese avanzado más en el dominio de la astronomía. Pero ninguno de los matemáticos que trabajaron en el

[21] *Nachrichten König. Ges. der Wiss. zu Gött.*, 15, 1868, 193-221 = *Wiss. Abh.*, 2, 618-639.

[22] *Theorie der Transformationsgruppen*, 3, 437-543, 1893.

período siguiente creyó que estas geometrías no euclídeas básicas tendrían trascendencia física. Cayley, Klein y Poincaré, aunque consideraron este asunto, afirmaron que nunca necesitaríamos mejorar o abandonar la geometría euclídea. El modelo de la pseudoesfera de Beltrami había hecho real a la geometría no euclídea en un sentido matemático (aunque no físico) porque daba una interpretación fácilmente visualizable de la geometría de Lobatchevsky, pero al precio de cambiar la recta del borde de la regla por las geodésicas y la pseudoesfera. Análogamente los modelos de Beltrami-Klein y Poincaré dieron sentido a la geometría no euclídea cambiando los conceptos ya sea de recta, distancia o medida de ángulos, o los tres al mismo tiempo, y representándolos en el espacio euclídeo. Pero el pensamiento de que el espacio físico podría ser no euclídeo bajo la interpretación usual de línea recta o bajo alguna otra interpretación fue rechazado. De hecho, la mayoría de los matemáticos trataban a la geometría no euclídea como una curiosidad lógica.

Cayley fue un firme partidario del espacio euclídeo y aceptó las geometrías no euclídeas solamente en cuanto que podían ser realizadas en el espacio euclídeo por medio de nuevas fórmulas para la distancia. En 1883, en su discurso presidencial a la «British Association for the Advancement of Science» (Asociación Británica para el Progreso de la Ciencia),[23] dijo que los espacios no euclídeos eran *a priori* una idea equivocada, pero que las geometrías no euclídeas eran aceptables porque resultaban meramente de un cambio en la función distancia del espacio euclídeo. No garantizó la existencia independiente de las geometrías no euclidianas sino que las trató como una clase de estructuras euclídeas especiales, o como una manera de representar las relaciones proyectivas en la geometría euclídea. Su opinión era que

> El duodécimo [décimo] axioma de Euclides en la forma que dio Playfair no necesita demostración, sino que es parte de nuestra noción de espacio, del espacio físico de nuestra propia experiencia —esto es, el espacio con el que llegamos a conocer por medio de la experiencia, pero que es la representación subyacente en el fundamento de toda experiencia exterior.
>
> Se puede decir que la opinión de Riemann es que, teniendo *in intellectu* una noción más general de espacio (de hecho una noción de espacio no euclídeo), aprendemos por medio de la experiencia que el espacio (el espacio

[23] *Collected Math. Papers*, 11, 429-459.

físico de la experiencia) es, si no exactamente, al menos en el más alto grado de aproximación, espacio euclídeo.

Klein consideró el espacio euclídeo como el espacio fundamental necesario. Las otras geometrías eran meramente euclídeas con las nuevas funciones para las distancias. Las geometrías no euclídeas estaban en efecto subordinadas a la geometría euclídea.

El juicio de Poincaré fue más liberal. La ciencia siempre debería tratar de usar la geometría euclídea y variar las leyes de la física donde fuese necesario. La geometría euclídea puede no ser verdadera pero es la más conveniente. Una geometría no puede ser más verdadera que otra; solamente puede ser más conveniente. El hombre crea la geometría y después adapta las leyes físicas a ella para hacer que la geometría y las leyes físicas coincidan con el mundo. Poincaré insistió [24] en que, aun si la suma de los ángulos de un triángulo resultara ser mayor de 180°, sería mejor suponer que la geometría euclídea describe el espacio físico y que la luz viaja siguiendo curvas porque la geometría euclídea es más sencilla. Por supuesto que los acontecimientos probaron que estaba equivocado. No es sólo la simplicidad de la geometría lo que cuenta para la ciencia sino la simplicidad de la teoría científica entera. Claramente, los científicos del siglo XIX estaban aún atados a la tradición por sus nociones acerca de lo que tiene sentido físico. El advenimiento de la teoría de la relatividad forzó un cambio drástico en la actitud hacia la geometría no euclídea.

La ilusión de los matemáticos en el sentido de que aquello en lo que están trabajando en el momento es el asunto más importante concebible se ilustra nuevamente por su actitud hacia la geometría proyectiva. La obra que hemos examinado en este capítulo muestra ciertamente que la geometría proyectiva es fundamental para muchas geometrías. Sin embargo, no abarca a la geometría riemanniana, evidentemente vital, y al cuerpo creciente de la geometría algebraica. De cualquier modo, Cayley afirmó en su artículo de 1859 (sec. 3) que «La geometría proyectiva es toda la geometría y recíprocamen-

---

[24] *Bull. Soc. Math. de France*, 15, 1887, 203-216 = *Œuvres*, 11, 79-91. Expresó esta opinión nuevamente en un artículo «Les Géométries non-euclidiennes» en la *Revue Génerale des Sciences*, 2, 1891, n. 23. Hay una traducción inglesa en *Nature* 45, 1892, 404-407. Véase también su *Science and Hypothesis*, cap. 3, en *The Foundations of Science*, the Science Press, 1946.

te.»[25] Bertrand Russell, en su *Essay on the Foundations of Geometry* (Ensayo sobre los fundamentos de la geometría) (1897), también creía que la geometría proyectiva era necesariamente la forma *a priori* de cualquier geometría del espacio físico. Hermann Hankel, a pesar de la atención que dio a la historia,[26] no vaciló en decir en 1869 que la geometría proyectiva es el camino real hacia todas las matemáticas. Nuestro examen de la evolución ya realizada muestra claramente que los matemáticos pueden fácilmente verse desorientados al dejarse llevar por sus entusiasmos.

Bibliografía

Beltrami, Eugenio: *Opere matematiche,* Ulrico Hoepli, 1902, vol. 1.

Bonola, Roberto: *Non-Euclidean Geometry,* Dover (reimp.), 1955, 129-264.

Coolidge, Julian L.: *A History of Geometrical Methods,* Dover (reimp.), 1963, 68-87.

Klein, Felix: *Gesammelte mathematische Abhandlungen,* Julius Springer, 1921-1923, vols. 1 y 2.

Pasch, Moritz, y Max Dehn: *Vorlesungen über neuere Geometrie,* 2.ª ed., Julius Springer, 1926, 185-239. Hay versión castellana, M. Pasch, *Lecciones de geometría moderna,* Madrid, Junta de Ampliación de Estudios, 1912.

Pierpont, James: «Non-Euclidean Geometry. A Retrospect», *Amer. Math. Soc. Bull.,* 36, 1930, 66-76.

Russell, Bertrand: *An Essay on the Foundations of Geometry* (1897), Dover (reimp.), 1956.

---

[25] Cayley usó el término «geometría descriptiva» para la geometría proyectiva.

[26] *Die Entwicklung der Mathematik in den letzten Jahrhunderten* (La Evolución de las Matemáticas durante los últimos siglos), 1869; 2.ª ed., 1884.

Capítulo 39
LA GEOMETRIA ALGEBRAICA

> En estos días el ángel de la topología y el demonio del álgebra abstracta luchan por el alma de cada dominio de las matemáticas.
>
> HERMANN WEYL

1. Antecedentes

Mientras se estaban creando las geometrías no euclídea y riemanniana, los investigadores de la geometría proyectiva estaban dedicados a su propio tema. Como hemos visto, las dos áreas fueron ligadas por la obra de Cayley y Klein. Después de que el método algebraico llegase a ser ampliamente usado en la geometría proyectiva, el problema de reconocer qué propiedades de las figuras geométricas eran independientes de la representación en coordenadas pedía atención, y esto motivó el estudio de los invariantes algebraicos.

Las propiedades proyectivas de las figuras geométricas son aquellas que quedan invariantes bajo transformaciones lineales de las figuras. Mientras los matemáticos estudiaban estas propiedades, se permitieron ocasionalmente considerar transformaciones de mayor grado y buscar aquellas propiedades de las curvas y superficies que son invariantes bajo estas últimas transformaciones. La clase de transformaciones, que pronto reemplazaron a las transformaciones lineales como interés preferido, es llamada birracional porque éstas se expresan algebraicamente como funciones racionales de las coordenadas y las transformaciones inversas también son funciones racionales

de sus coordenadas. La concentración en las transformaciones birracionales resultó indudablemente del hecho de que Riemann las había usado en su trabajo sobre integrales y funciones abelianas y de hecho, como veremos, los primeros grandes pasos en el estudio de las transformaciones birracionales de las curvas fueron guiados por lo que había hecho Riemann. Estos dos temas formaron el contenido de la geometría algebraica durante la última parte del siglo XIX.

El término geometría algebraica es inadecuado, ya que originalmente se refería a toda la obra desde la época de Fermat y Descartes en la que el álgebra había sido aplicada a la geometría; en la última parte del siglo XIX se aplicó al estudio de los invariantes algebraicos y las transformaciones birracionales. En el siglo XX se refiere al último campo mencionado.

2. La teoría de invariantes algebraicos

Como ya hemos observado, la determinación de las propiedades geométricas de las figuras que se representan y estudian por medio de la representación coordenada exige el discernimiento de aquellas expresiones algebraicas que permanecen invariantes bajo los cambios de coordenadas. Visto en forma alternativa, la transformación proyectiva de una figura en otra por medio de una transformación lineal conserva algunas propiedades de la figura. Los invariantes algebraicos representan estas propiedades geométricas invariantes.

El tema de los invariantes algebraicos ya había surgido previamente en la teoría de números (cap. 34, sec. 5) y particularmente en el estudio de cómo se transforman las formas cuadráticas binarias.

$$f = ax^2 + 2bxy + xy^2 \qquad (1)$$

cuando x y y son transformadas por la transformación lineal T, a saber,

$$x = \alpha x' + \beta y', \qquad y = \gamma x' + \delta y', \qquad (2)$$

donde $\alpha\delta - \beta\gamma = r$. La aplicación de T a f da

$$f' = a'x'^2 + 2b'x'y' + c'y'^2. \qquad (3)$$

En la teoría de números las cantidades a, b, c, α, β, γ y δ son enteros y $r = 1$. Sin embargo, es cierto generalmente que el discriminante D de f satisface la relación

$$D' = r^2 D. \qquad (4)$$

Las transformaciones lineales de la geometría proyectiva son más generales porque los coeficientes de las formas y las transformaciones no están restringidos a ser enteros. El término invariantes algebraicos se usa para distinguir aquellos que surgen bajo estas transformaciones lineales más generales de los invariantes modulares de la teoría de números y, a estos efectos, de los invariantes diferenciales de la geometría riemanniana.

Para discutir la historia de la teoría de los invariantes algebraicos necesitamos algunas definiciones. La forma de n-ésimo grado en una variable

$$f(x) = a_o x^n + a_1 x^{n-ı} + \ldots + a_n$$

se convierte en coordenadas homogéneas en la forma binaria

$$f(x_1,x_2) + a_o x_1^n + a\dagger q, x_1^{n-1} x_2 + \ldots + a_n x_2^n. \qquad (5)$$

En tres variables las formas son llamadas ternarias; en cuatro variables, cuaternarias, etc. Las definiciones que siguen se aplican a las formas en n variables.

Supóngase que sometemos la forma binaria a una transformación T de la forma (2). Mediante T, la forma $f(x_1,x_2)$ es transformada en la forma

$$F(X_1,X_2) = A_o X_1^n + A_1 X_1^{n-1} X_2 + A_n X_2^n.$$

Los coeficientes de F diferirán de los de f y las raíces de $F = 0$ diferirán de las de $f = 0$. Cualquier función I de los coeficientes de f que satisface la relación

$$I(A_o, A_1, \ldots, A_n) = r^w I(a_o, a_1, \ldots, a_n)$$

es llamada un invariante de f. Si $w = 0$, el invariante se llama invariante absoluto de f. El grado del invariante es el grado en los coe-

ficientes y el peso es w. El discriminante de una forma binaria es un invariante, como lo ejemplifica (4). En este caso el grado es 2 y el peso es 2. La significación del discriminante de cualquier ecuación polinomial $f(x) = 0$ es que su anulación es la condición de que $f(x) = 0$ tenga raíces iguales o, geométricamente, que el lugar geométrico de $f(x) = 0$, el cual está formado por una serie de puntos, tengan dos puntos que coincidan. Esta propiedad es claramente independiente del sistema de coordenadas.

Si dos formas binarias (o más)

$$f_1 = a_0 x_1^m + \dots + a_m x_2^m$$
$$f_2 = b_0 x_1^n + \dots + b_n x_2^n$$

son transformadas por T en

$$F_1 = A_0 X_1^m + \dots + A_m X_2^m$$
$$F_2 = B_0 X_1^n + \dots + B_n X_2^n$$

entonces cualquier función I de los coeficientes que satisfaga la relación

$$I(A_0, \dots, A_m, B_0, \dots, B_n) = r^w I(a_0, \dots, a_m, b_0, \dots, b_n) \qquad (6)$$

se dice que es un invariante simultáneo de las dos formas. Así, las formas lineales $a_1x_1 + b_1x_2$ y $a_2x_1 + b_2x_2$ tienen como invariante simultáneo la resultante $a_1b_2 - a_2b_1$ de las dos formas. Geométricamente, la anulación de la resultante significa que las dos formas representan al mismo punto (en coordenadas homogéneas). Dos formas cuadráticas $f_1 = a_1x_1^2 + 2b_1x_1x_2 + c_1x_2^2$ y $f_2 = a_2x_1^2 + 2b_2x_1x_2 + c_2x_2^2$ poseen un invariante simultáneo

$$D_{12} = a_1c_2 - 2b_1b_2 + a_2c_1$$

cuya anulación expresa el hecho de que f_1 y f_2 representan pares armónicos de puntos.

Más allá de los invariantes de una forma o de un sistema de formas hay covariantes. Cualquier función C de los coeficientes *y* variables de f que es un invariante por T excepto para una potencia del módulo (determinante) de T es llamada un covariante de f. Así, para las formas binarias, un covariante satisface la relación

$$C(A_0, A_1, \ldots, A_n, X_1, X_2) = r^{\omega}C(a_0, a_1, \ldots, a_n, x_1, x_2).$$

Las definiciones de covariantes absolutos y simultáneos son análogas a las de los invariantes. El grado de un covariante en los coeficientes es llamado su grado y el grado en sus variables es llamado orden. Los invariantes son, por tanto, covariantes de orden cero. Sin embargo, algunas veces la palabra invariante se usa para referirse a un invariante en el sentido más estricto o a un covariante.

Un covariante de f representa alguna figura que no solamente está relacionada con f, sino relacionada proyectivamente. Así, el jacobiano de dos formas binarias cuadráticas $f(x_1,x_2)$ y $\phi(x_1,x_2)$, a saber

$$\begin{vmatrix} \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} \\ \dfrac{\partial \phi}{\partial x_1} & \dfrac{\partial \phi}{\partial x_2} \end{vmatrix}$$

es un covariante simultáneo de peso 1 de las dos formas. Geométricamente, el conjunto donde el jacobiano es igual a cero representa un par de puntos que es armónico con cada uno de los pares originales representados por f y ϕ y la propiedad armónica es proyectiva.

El hessiano de una forma binaria introducido por Hesse [1],

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{vmatrix}$$

es un covariante de peso 2. Su significado geométrico es demasiado complicado para concederle espacio aquí (*Cf.* cap. 35, sec. 5). El concepto de hessiano y su covariancia se aplica a cualquier forma en n variables.

El trabajo sobre invariantes algebraicos fue iniciado en 1841 por George Boole (1815-1864), cuyos resultados [2] eran limitados. Lo que tiene mayor relevancia es que Cayley fue atraído hacia este tema

[1] *Joür. fur Math.*, 28, 1844, 68-96 = *Ges. Abh.*, 89-122.

[2] *Cambridge Mathematical Journal*, 3, 1841, 1-20; y 3, 1842, 106-119.

por la obra de Boole y también hizo que Sylvester se interesara en el asunto. A ellos se les unió George Salmon (1819-1904), quien fue profesor de matemáticas en el Trinity College en Dublín desde 1840 hasta 1866, y después se convirtió en profesor de teología en esa institución. Estos tres autores realizaron tanto trabajo sobre los invariantes que en una de sus cartas Hermite los apodó la trinidad invariante.

En 1841 Cayley empezó a publicar artículos matemáticos sobre el aspecto algebraico de la geometría proyectiva. El artículo de Boole de 1841 sugirió a Cayley el cálculo de los invariantes de las funciones homogéneas de grado *n*. Llamó a los invariantes derivados y después hiperdeterminantes; el término invariante se le debe a Sylvester.[3] Cayley, empleando las ideas de Hesse y de Eisenstein sobre determinantes, desarrolló una técnica para generar sus «derivados». Después publicó diez artículos sobre cuánticos en las *Philosophical Transactions*, desde 1854 hasta 1878.[4] Cuánticos fue el término que adoptó para los polinomios homogéneos en 2, 3 o más variables. Cayley llegó a interesarse tanto en los invariantes que los investigó por el propio interés que presentaban. También inventó un método simbólico para tratar los invariantes.

En el caso particular de la forma cuártica binaria

$$f = ax_1^4 + 4bx_1^3x_2 + 6cx_1^2x_2^2 + 4dx_1x_2^3 + ex_2^4$$

Cayley mostró que el hessiano H y el jacobiano de f y H son covariantes y que

$$g_2 = ae - 4bd + 3c^2$$

$$g_3 = \begin{vmatrix} a & b & c \\ b & c & d \\ c & d & e \end{vmatrix}$$

son invariantes. A estos resultados agregaron muchos más Sylvester y Salmon.

Otro personaje que aportó resultados, Ferdinand Eisenstein,

---

[3] *Coll. Math. Papers*, I, 273.
[4] *Coll. Math. Papers*, 2, 4, 6, 7, 10.

quien estaba más interesado en la teoría de números, ya había encontrado para la forma cúbica binaria [5]

$$f = ax_1^3 + 3bx_1^2x_2 + 3cx_1x_2^2 + dx_2^3$$

que el covariante más sencillo de segundo grado es su hessiano H y el invariante más sencillo es

$$3b^2 c^2 + 6abcd - 4b^3d - 4ac^3 - a^2d^2,$$

el cual es el determinante del hessiano cuadrático, al igual que el discriminante de f. También el jacobiano de f y H es otro covariante de orden tres. Después Siegfried Heinrich Aronhold (1819-1884), quien empezó a trabajar con invariantes en 1849, proporcionó invariantes para las formas cúbicas ternarias.[6]

El primer problema importante que afrontaron los fundadores de la teoría de invariantes fue el descubrimiento de invariantes particulares. Esta fue la dirección de la investigación desde aproximadamente 1840 hasta 1870. Como podemos ver, se pueden construir muchas de estas funciones porque algunos invariantes tales como el jacobiano y el hessiano son ellos mismos formas que tienen invariantes y porque algunos invariantes, tomados junto con la forma original, dan un nuevo sistema de formas que tienen invariantes simultáneos. Docenas de matemáticos sobresalientes, incluyendo a los pocos que ya hemos mencionado, calcularon invariantes particulares.

La continuación del cálculo de invariantes llevó al problema principal de la teoría de invariantes, el cual surgió después de que se encontraran muchos invariantes especiales o particulares; éste consistía en encontrar un sistema completo de invariantes. Lo que esto quiere decir es encontrar para una forma de un número de variables y grado dados, el número más pequeño posible de invariantes enteros racionales o covariantes tal que cualquier otro invariante o covariante entero racional pudiera expresarse como una función entera racional con coeficientes numéricos de este conjunto completo. Cayley demostró que los invariantes y covariantes encontrados por Eisenstein para la forma cúbica binaria y los que obtuvo para la forma cuártica binaria son un sistema completo para los casos respectivos.[7]

---

[5] *Jour. für Math.*, 27, 1844, 89-106, 319-321.

[6] *Jour. für Math.*, 55, 1858, 97-191; y 62, 1863, 281-345.

[7] *Phil. Trans.*, 146, 1856, 101-126 = *Coll. Math. Papers*, 2, 250-275.

Esto dejó abierto el problema de un sistema completo para otras formas.

La existencia de un sistema completo finito o base para las formas binarias de cualquier grado dado fue establecida primero por Paul Gordan (1837-1912), quien dedicó la mayor parte de su vida a este tema. Su resultado [8] es que a cada forma binaria $f(x_1,x_2)$ le corresponde un sistema completo finito de invariantes y covariantes enteros racionales. Gordan contó con el auxilio de teoremas debidos a Clebsch y el resultado es conocido como el teorema de Clebsch-Gordan. La demostración es larga y difícil. Gordan también demostró [9] que cualquier *sistema* finito de formas binarias tiene un sistema completo finito de invariantes y covariantes. Las pruebas de Gordan mostraron cómo calcular los sistemas completos.

Se obtuvieron varias extensiones limitadas de los resultados de Gordan durante los veinte años siguientes. Gordan mismo proporcionó el sistema completo para la forma cuadrática ternaria,[10] para la forma cúbica ternaria,[11] y para un sistema de dos y tres cuadráticas ternarias.[12] Para la cuártica ternaria especial $x_1^3x_2 + x_2^3x_3 + x_3^3x_1$, Gordan dio un sistema completo de 54 formas fundamentales.[13]

En 1886, Franz Mertens (1840-1927) [14] volvió a demostrar el teorema de Gordan para los sistemas binarios por un método inductivo. Supuso que el teorema era cierto para cualquier conjunto dado de formas binarias y después demostró que debería seguir siendo cierto cuando se aumentaba en uno el grado de una de las formas. No mostró explícitamente el conjunto finito de invariantes y covariantes independientes pero sí demostró que existía. El caso más sencillo, una forma lineal, fue el punto de partida de la inducción y tal forma tiene como covariantes solamente potencias de sí misma.

Hilbert, en 1885, después de escribir una tesis doctoral sobre invariantes, [15] también volvió a demostrar en 1888 [16] el teorema de

---

[8] *Jour. für Math.*, 69, 1868, 323-354.
[9] *Math. Ann.*, 2, 1870, 227-280.
[10] R. Clebsch y F. Lindemann, *Vorlesungen über Geometrie*, I, 1876, p. 291.
[11] *Math. Ann.*, 1, 1869, 56-89, 90-128.
[12] Clebsch-Lindemann, p. 288.
[13] *Math. Ann.*, 17, 1880, 217-233.
[14] *Jour. für Math.*, 100, 1887, 223-230.
[15] *Math. Ann.*, 30, 1887, 15-29 = *Ges. Abh.*, 2, 102-116.
[16] *Math. Ann.*, 33, 1889, 223-226 = *Ges. Abh.*, 2, 162-164.

Gordan de que cualquier sistema dado de formas binarias tiene un sistema de invariantes y covariantes completo finito. Su demostración fue una modificación de la de Mertens. Ambas pruebas eran mucho más fáciles que la de Gordan. Pero la prueba de Hilbert tampoco presentó un proceso para encontrar el sistema completo.

En 1888, Hilbert sorprendió a la comunidad matemática anunciando un enfoque totalmente nuevo del problema de demostrar que cualquier forma de grado dado y número de variables dado, y que cualquier sistema de formas dado en cualquier número de variables dado, tiene un sistema completo finito de invariantes y covariantes enteros racionales independientes.[17] La idea básica de este nuevo enfoque consistía en olvidarse de los invariantes por el momento y considerar la cuestión de si, dado un sistema infinito de expresiones enteras racionales en un número finito de variables, ¿bajo qué condiciones existe un número finito de estas expresiones, una base, en términos de la cual todas las demás sean expresables como combinaciones lineales con funciones enteras racionales de las mismas variables como coeficientes? La respuesta es: siempre. Más específicamente, el teorema de la base de Hilbert, que precede al resultado sobre invariantes, es como sigue: por forma algebraica entendemos una función homogénea entera racional en n variables con coeficientes en algún dominio de racionalidad definido (cuerpo). Dada una colección de infinitas formas de cualquier grado en n variables, entonces existe un número finito (una base) F_1, F_2, ..., F_m tal que cualquier forma F de la colección puede escribirse como

$$F = A_1F_1 + A_1F_2 + \ldots + A_mF_m$$

donde A_1, A_2, ..., A_m son formas apropiadas en las n variables (no necesariamente en el sistema infinito) con coeficientes en el mismo dominio que los coeficientes del sistema infinito.

En la aplicación de este teorema a los invariantes y covariantes, el resultado de Hilbert establece que para cualquier forma o sistema de formas existe un número finito de invariantes y covariantes enteros racionales por medio de los cuales se pueden expresar todos los demás invariantes o covariantes enteros racionales como combi-

[17] *Math. Ann.*, 36, 1890, 473-534 = *Ges. Abh.*, 2, 199-257, y en artículos subsiguientes hasta 1893.

nación lineal de los del conjunto finito. Esta colección finita de invariantes y covariantes es el sistema completo de invariantes.

La demostración de existencia de Hilbert era tan sencilla con respecto al cálculo laborioso de Gordan, que éste no pudo evitar exclamar: «Esto no son matemáticas; es teología». Sin embargo, reconsideró el asunto y posteriormente dijo: «Me he convencido a mí mismo de que la teología también tiene sus ventajas.» De hecho, él mismo simplificó la demostración de existencia dada por Hilbert.[18]

En los años 1880 y 1890 se pensó en la teoría de los invariantes como unificadora de muchas áreas de las matemáticas. Esta teoría fue el «álgebra moderna» del período. Sylvester dijo en 1864:[19] «De la misma manera que todos los caminos conducen a Roma, así encuentro, en mi propio caso al menos, que todas las investigaciones algebraicas, tarde o temprano, van a dar al Capitolio del álgebra moderna sobre cuyo pórtico brillante se inscribe la Teoría de los Invariantes.» Pronto la teoría se convirtió en un fin en sí misma, independientemente de sus orígenes en la teoría de números y en la geometría proyectiva. Los que se dedicaban al estudio de los invariantes algebraicos persistieron en demostrar toda clase de identidades algebraicas, tuvieran o no significación geométrica. Maxwell, cuando era estudiante en Cambridge, dijo que algunos hombres veían al universo entero en términos de quínticas y cuánticas.

Por otro lado, los físicos de finales del siglo XIX no tomaron cartas en el asunto. Ciertamente Tait observó una vez acerca de Cayley, «¿No es una vergüenza que ese hombre tan sobresaliente dedique sus habilidades a tales cuestiones totalmente inútiles?» Sin embargo, el asunto sí tuvo impacto en la física, indirecta y directamente, en gran medida a través del trabajo en invariantes diferenciales.

A pesar del enorme entusiasmo por la teoría de invariantes en la segunda mitad del siglo XIX, el tema perdió atractivo tal y como fue concebido e investigado durante ese período. Los matemáticos dicen que Hilbert mató la teoría de invariantes porque había resuelto todos los problemas. Hilbert le escribió a Minkowski en 1893 que ya no se dedicaría más a ese tema, y dijo en un artículo de 1893 que las metas generales más importantes de la teoría ya se habían alcanzado. Sin embargo, estaba lejos de ser así. El teorema de Hilbert no mostraba cómo calcular los invariantes para cualquier forma o sis-

---

[18] *Nachrischten König. Ges. der Wiss. zu Gött.*, 1899, 240-242.

[19] *Phil. Trans.*, 154, 1864, 579-666 = *Coll. Math. Papers*, 2, 376-479, p. 380.

tema de formas dados, y por tanto no podía proporcionar ni siquiera un solo invariante significativo. La búsqueda de invariantes específicos que tuvieran significado geométrico o físico era aún importante. Incluso el cálculo de una base para las formas de grado y número de variables dados podría ser valioso.

Lo que «mató» a la teoría de invariantes en el sentido que se entendía el tema en el siglo XIX es la usual colección de factores que mataron muchas otras actividades investigadas con demasiado entusiasmo. Los matemáticos siguen a los maestros. El pronunciamiento de Hilbert y el hecho de que él mismo abandonó el campo, ejercieron gran influencia en otros. También el cálculo de invariantes específicos significantes había llegado a ser más difícil después de que se lograran los resultados más fácilmente obtenibles.

El cálculo de invariantes algebraicos no acabó con la obra de Hilbert. Emmy Noether (1882-1935), una alumna de Gordan, escribió su tesis doctoral en 1907 «Sobre Sistemas Completos de Invariantes para las Formas Bicuadráticas Ternarias».[20] También dio un sistema completo de formas covariantes para una cuártica ternaria, 331 en total. En 1910 extendió el resultado de Gordan a *n* variables.[21]

La historia posterior de la teoría algebraica de invariantes pertenece al álgebra abstracta moderna. La metodología de Hilbert llevó a primer plano la teoría abstracta de los módulos, anillos y cuerpos. Hilbert demostró, en este lenguaje, que todo sistema modular (un ideal en la clase de los polinomios en *n* variables) tiene una base que consiste en un número finito de polinomios, o que todo ideal en un dominio polinomial de *n* variables posee una base finita siempre que en el dominio de los coeficientes de los polinomios todo ideal tenga una base finita. De 1911 a 1919, Emmy Noether escribió muchos artículos sobre bases finitas para varios casos distintos usando la técnica de Hilbert y la suya propia. En su desarrollo del siglo XX dominó el punto de vista algebraico abstracto. Tal y como se quejó Eduard Study en su texto sobre la teoría de invariantes, hubo falta de interés por problemas específicos y solamente se investigaron los métodos abstractos.

---

[20] *Jour. für Math.*, 134, 1908, 23-90.
[21] *Jour. für Math.*, 139, 1911, 118-154.

3. El concepto de transformación birracional

Vimos en el capítulo 35 que, especialmente durante las décadas tercera y cuarta del siglo XIX, el trabajo en geometría proyectiva se dirigió hacia las curvas de grado superior. Sin embargo, antes de que este trabajo hubiese avanzado demasiado, hubo un cambio en la naturaleza del estudio. El punto de vista proyectivo significa transformaciones lineales en coordenadas homogéneas. Gradualmente entraron en escena transformaciones de segundo grado (y superior) y el interés se desvió hacia las transformaciones birracionales. Tales transformaciones, para el caso de dos coordenadas no homogéneas, son de la forma

$$x' = \phi(x, y), \qquad y' = \psi(x, y)$$

donde ϕ y ψ son funciones racionales en x e y, y además x e y pueden expresarse como funciones racionales de x' e y'. En coordenadas homogéneas x_1, x_2 y x_3 las transformaciones son de la forma

$$x'_i = F_i\,(x_1, x_2, x_3), \qquad i = 1, 2, 3$$

y la inversa es

$$x_i = G_i(x_?, x_?, x_i), \qquad i = 1, 2, 3$$

donde F_i y G_i son polinomios homogéneos de grado n en sus respectivas variables. La correspondencia es uno-a-uno, excepto que cada uno de los puntos de un conjunto finito de ellos puede corresponder a una curva.

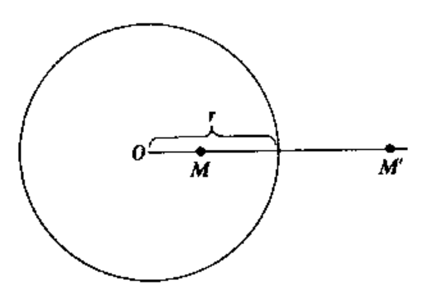


FIGURA 39.1

Como ejemplo de una transformación birracional tenemos la inversión con respecto a un círculo. Geométricamente, esta transformación (fig. 39.1) va de M a M' o de M' a M por medio de la ecuación que la define

$$OM \cdot OM' = r^2,$$

donde r es el radio del círculo. Algebraicamente, si fijamos un sistema de coordenadas en O, por el teorema de Pitágoras se llega a que

$$x' = \frac{r^2x}{x^2 + y^2}, \qquad y' = \frac{r^2y}{x^2 + y^2}, \tag{7}$$

donde M es (x, y) y M' es (x', y'). Con esta transformación los círculos se transforman en círculos o líneas rectas y viceversa. La inversión es una transformación que lleva todo el plano en sí mismo y tales transformaciones birracionales son llamadas transformaciones de Cremona. Otro ejemplo de una transformación de Cremona en tres variables (homogéneas) es la transformación cuadrática

$$x_i = x_2x_3, \qquad x_2 = x_3x_1, \qquad x_3 = x_1x_2 \tag{8}$$

cuya inversa es

$$x_1 = x_2x_3, \qquad x_2 = x_3x_1, \qquad x_3 = x_1x_2.$$

El término transformación birracional también se usa en un sentido más general, a saber, donde la transformación de los puntos de una curva en los de otra es birracional pero la transformación no necesita ser birracional en todo el plano. Así (en coordenadas no homogéneas) la transformación

$$X = x^2, \qquad Y = y \tag{9}$$

no es inyectiva en todo el plano pero sí lleva cualquier curva C situada a la derecha del eje y en otra por una correspondencia uno-a-uno.

La transformación de inversión fue la primera de las transformaciones birracionales que aparecieron. En situaciones limitadas fue usada por Poncelet en su *Traité* de 1822 (π370) y después por Plüc-

ker, Steiner, Quetelet y Ludwig Immanuel Magnus (1790-1861). Fue estudiada extensamente por Möbius [22] y su uso en física fue reconocido por Lord Kelvin [23] y por Liouville,[24] quien la llamó transformación por radios recíprocos.

En 1854, Luigi Cremona (1830-1903), que fue profesor de matemáticas en varias universidades italianas, introdujo la transformación birracional general (del plano completo en sí mismo) y escribió importantes artículos sobre ello.[25] Max Noether (1844-1921), padre de Emmy Noether, demostró entonces el resultado fundamental [26] de que una transformación plana de Cremona puede construirse a partir de una sucesión de transformaciones cuadráticas y lineales. Jacob Rosanes (1842-1922) encontró este resultado independientemente [27] y demostró también que todas las transformaciones algebraicas uno-a-uno del plano deben ser transformaciones de Cremona. Las demostraciones de Noether y Rosanes fueron completadas por Guido Castelnuovo (1865-1952).[28]

4. El enfoque de la teoría de funciones en geometría algebraica

Aunque era clara la naturaleza de las transformaciones birracionales, el desarrollo del área de la geometría algebraica como el estudio de los invariantes bajo tales transformaciones fue, al menos en el siglo XIX, insatisfactorio. Se emplearon varios enfoques; los resultados eran fragmentarios y sin relación; la mayoría de las demostraciones estaban incompletas; y se obtuvieron muy pocos teoremas importantes. La variedad de los enfoques ha dado lugar a marcadas diferencias en los lenguajes usados. Las metas del área también eran vagas. Aunque el tema principal había sido la invariancia bajo transformaciones birracionales, la materia cubre la investigación de las propiedades de las curvas, superficies y estructuras de mayores di-

[22] *Theorie der Kreisverwandschaft* (Teoría de la Inversión), *Abh. Konig. Sach. Ges. der Wiss.*, 2, 1855, 529-565 = *Werke*, 2, 243-345.

[23] *Jour. de Math.*, 10, 1845, 364-367.

[24] *Jour. de Math.*, 12, 1847, 265-290.

[25] *Gior. di Mat.*, 1, 1863, 305-311 = *Opere*, 1, 54-61; y 3, 1865, 269-280, 363-376 = *Opere*, 2, 193-218.

[26] *Math. Ann.*, 3, 1871, 165-227, p. 167 en particular.

[27] *Jour. für Math.*, 73, 1871, 97-110.

[28] *Atti Accad. Torino*, 36, 1901, 861-874.

mensiones. Desde este punto de vista no hay muchos resultados importantes. Daremos unos pocos ejemplos de lo que se hizo.

Clebsch planteó uno de los primeros enfoques. (Rudolf Friedrich) Alfred Clebsch (1833-1872) estudió bajo el asesoramiento de Hesse en Königsberg de 1850 a 1854. En la primera etapa de su trabajo estuvo interesado en la física matemática y de 1858 a 1863 fue profesor de mecánica teórica en Karlsruhe y después profesor de matemáticas en Giessin y Göttingen. Trabajó sobre problemas dejados por Jacobi en el cálculo de variaciones y la teoría de las ecuaciones diferenciales. En 1862 publicó el *Lehrbuch der Elasticität (Tratado de Elasticidad)*. Sin embargo, su labor principal la llevó a cabo en invariantes algebraicos y geometría algebraica.

Clebsch había trabajado sobre las propiedades proyectivas de las curvas y superficies de tercero y cuatro grado hasta aproximadamente 1860. Conoció a Paul Gordan en 1863 y supo del trabajo de Riemann en la teoría de las funciones de variable compleja. Clebsch hizo entonces que esta teoría sirviera a la teoría de las curvas.[29] Este enfoque es llamado trascendente. A pesar de que Clebsch llevó a cabo la conexión entre las funciones complejas y las curvas algebraicas, en una carta que envió a Gustav Roch admitió que no podía entender la obra de Riemann sobre funciones abelianas ni las contribuciones de Roch en su disertación.

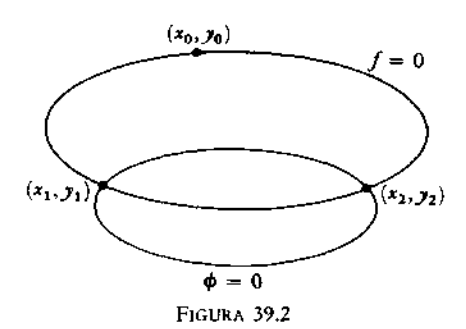


FIGURA 39.2

Clebsch reinterpretó la teoría de las funciones complejas de la siguiente manera: la función $f(w,z) = 0$, donde z y w son variables complejas, geométricamente requiere de una superficie de Riemann para z y un plano o una parte de un plano para w o, si se prefiere,

[29] *Jour. für Math.*, 63, 1864, 189-243.

de una superficie de Riemann a la cual se asigna un par de valores z y w a cada uno de sus puntos. Considerando sólo la parte real de z y w, la ecuación $f(w,z) = 0$ representa una curva en el plano cartesiano real. Es posible que z y w tengan aún valores complejos que satisfagan $f(w,z) = 0$, pero éstos no se dibujan. Este punto de vista de las curvas reales con puntos complejos era ya familiar a partir del trabajo en geometría proyectiva. A la teoría de las transformaciones birracionales de las superficies corresponde una teoría de las transformaciones birracionales de las curvas planas. Con la reinterpretación ahora descrita, los puntos de ramificación de la superficie de Riemann corresponden a aquellos puntos de la curva donde una línea $x =$ const., corta a la curva en dos o más puntos consecutivos, esto es, o es tangente a la curva o pasa por una cúspide. Un punto doble de la curva corresponde a un punto en la superficie donde dos hojas se tocan exactamente, sin ninguna relación adicional. Los puntos múltiples de mayor orden en las curvas también corresponden a otras peculiaridades de las superficies de Riemann.

En lo que sigue utilizaremos las siguientes definiciones (*cf.* cap. 23, sec. 3): un punto múltiple (punto singular) P de orden $k > 1$ de una curva plana de grado n es un punto tal que una recta genérica que pasa por P corta a la curva en $n - k$ puntos. El punto múltiple es ordinario si las k tangentes en P son distintas. Al contar las intersecciones de una curva de n-ésimo grado y de otra de grado m-ésimo, se debe tener en cuenta la multiplicidad de los puntos múltiples sobre cada curva. Si es h en la curva C^n y k en C^m y si las tangentes en P de C^n son distintas de las de C^m, entonces el punto de intersección tiene multiplicidad hk. Se dice que una curva C' es adjunta de una curva C cuando los puntos múltiples de C son ordinarios o cúspides y si C' tiene un punto de orden de multiplicidad $k - 1$ en todo punto múltiple de C de orden k.

Clebsch [30] reformuló por primera vez el teorema de Abel (cap. 27, sec. 7) sobre integrales de primera clase en términos de curvas. Abel consideró una función racional fija $R(x,y)$ donde x y y están relacionadas por cualquier curva algebraica $f(x,y) = 0$, de modo que y sea una función de x. Supóngase (fig. 39.2) que $f = 0$ es cortada por otra curva algebraica

$$\phi(x, y, a_1, a_2, ..., a_k) = 0,$$

[30] *Jour. für Math.*, 63, 1864, 189-243.

donde los a_i son los coeficientes de $\phi = 0$. Sean (x_1,y_1), (x_2,y_2), ... (x_m,y_m) las intersecciones de $\phi = 0$ con $f = 0$. (El número m de ellas es el producto de los grados de f y ϕ.) Dado un punto (x_0,y_0) en $f = 0$, donde y_0 pertenece a una rama de $f = 0$, entonces podemos considerar la suma

$$I = \sum_{i=1}^{m} \int_{x_0,y_0}^{x_i,y_i} R(x, y)\, dx.$$

Los límites superiores x_i, y_i se encuentran en $\phi = 0$ y la integral I es una función de los límites superiores. Entonces hay un número característico p de estos límites que tienen que ser funciones algebraicas de los otros. Este número p depende solamente de f. Más aún, I puede expresarse como la suma de estas p integrales y de funciones racionales y logarítmicas de las x_i, y_i, $i = 1, 2, ..., m$. Además, si se hace variar la curva $\phi = 0$ variando los parámetros de a_1 hasta a_k, *entonces* x_i variará también en I se convertirá en una función de los a_i mediante los x_i. La función I de los a_i será racional en los a_i o, en el peor de los casos, involucrará funciones logarítmicas de las a_i.

Clebsch también aplicó a las curvas el concepto de Riemann de las integrales abelianas sobre superficies de Riemann, esto es, las integrales de la forma $\int g(x,y)dx$, donde g es una función racional y $f(x,y) = 0$. Para ilustrar las integrales de primera clase considérese una curva plana de cuarto grado C_4 sin puntos dobles. Aquí $p = 3$ y se tienen las tres integrales finitas en todas partes.

$$u_1 = \int \frac{x\, dx}{f_y}, \qquad u_2 = \int \frac{y\, dx}{f_y}, \qquad u_3 = \int \frac{dx}{f_y}.$$

Lo que se aplica a la C_4 también se cumple para curvas algebraicas arbitrarias $f(x,y) = 0$ de orden n. En lugar de las tres integrales finitas en todas partes hay entonces p de éstas integrales (donde p es el género de $f = 0$). Cada una de ellas tiene $2p$ módulos de periodicidad (cap. 27, sec. 8). Las integrales son de la forma

$$\int \frac{\phi(x, y)}{\partial f/\partial y}\, dx$$

donde ϕ es un polinomio (un adjunto) de grado precisamente $n - 3$, el cual se anula en los puntos dobles y cúspides de $f = 0$.

La contribución siguiente de Clebsch [31] consistió en introducir la noción de género como un concepto para clasificar las curvas. Si la curva tiene d puntos dobles entonces el género es $p = (1/2)(n-1)(n-2) - d$. Previamente ya se tenía la noción de la deficiencia de una curva (cap. 23, sec. 3), esto es, el máximo número posible de puntos dobles que podía poseer una curva de grado n, a saber $(n-1)(n-2)/2$, menos el número que realmente posee. Clebsch demostró [32] que para curvas con solamente puntos múltiples ordinarios (las tangentes son todas distintas) el género es igual a la deficiencia y es invariante por una transformación birracional de todo el plano en sí mismo.[33] La noción de género dada por Clebsch está relacionada con la conectividad de Riemann de una superficie de Riemann. La superficie de Riemann correspondiente a una curva de género p tiene conectividad $2p + 1$.

La noción de género puede utilizarse para establecer teoremas significativos acerca de las curvas. Jacob Lüroth (1844-1910) demostró [34] que una curva de género 0 puede ser transformada birracionalmente en una línea recta. Cuando el género es 1, Clebsch demostró que una curva puede transformarse birracionalmente en una curva de tercer grado.

Además de la clasificación de curvas por el género, Clebsch, siguiendo a Riemann, introdujo clases dentro de cada género. Riemann había considerado [35] la transformación birracional de sus superficies. Así, si $f(w,z) = 0$ es la ecuación de la superficie y si

$$w_1 = R_1(w,z),\ z_1 = R_2(w,z)$$

son funciones racionales, y si la transformación inversa es racional, entonces $f(w,z)$ puede transformarse en $F(w_1,z_1) = 0$. Dos ecuaciones algebraicas $F(w,z) = 0$ (o sus superfices) pueden transformarse birracionalmente una en la otra solamente si ambas tienen el mismo valor p. (El número de hojas no se conserva, en general). Para Rie-

---

31 *Jour. für Math.*, 64, 1865, 43-65.

32 *Jour. für Math.*, 64, 1865, 98-100.

33 Si los puntos múltiples (singulares) son de orden r_i entonces el género p de una curva C es $(n-1)(n-2)/2 - (1/2)\sum r_i(r_i - 1)$, donde la sumatoria se extiende sobre todos los puntos múltiples. El género es un concepto más refinado.

34 *Math. Ann.*, 9, 1876, 163-165.

35 *Jour. für Math.*, 54, 1857, 115-155 = *Werke*, 2.ª ed., 88-142.

mann no era necesaria ninguna demostración adicional. Estaba garantizada por la intuición.

Riemann (en el artículo de 1857) trató todas las ecuaciones (o las superficies) que son transformables birracionalmente una en la otra como pertenecientes a la misma clase. Tienen el mismo género p. Sin embargo, hay diferentes clases con el mismo valor p (porque los puntos de ramificación pueden diferir). La clase más general de género p es caracterizada por $3p - 3$ constantes (complejas) coeficientes en la ecuación cuando $p > 1$, por una constante cuando $p = 1$ y por cero constantes cuando $p = 0$. En el caso de funciones elípticas $p = 1$, y hay una constante. Las funciones trigonométricas, para las cuales $p = 0$, no tienen constante arbitraria alguna. El número de constantes fue llamado por Riemann el módulo de clase. Las constantes son invariantes por transformaciones birracionales. Clebsch, de manera parecida, puso a todas aquellas curvas que se pueden obtener de una dada por medio de una transformación birracional uno-a-uno en una clase. Las de una clase tienen necesariamente el mismo género, pero puede haber diferentes clases con el mismo género.

5. El problema de la uniformización

Clebsch volvió entonces su atención a lo que se llama el problema de la uniformización para curvas. Indiquemos primero exactamente a qué se refiere este problema. Dada la ecuación

$$w^2 + z^2 = 1 \tag{10}$$

la podemos representar en forma paramétrica como

$$z = sen\ t, \qquad w = cos\ t \tag{11}$$

o en la forma paramétrica

$$z = \frac{2t}{1+t^2}, \qquad w = \frac{1-t^2}{1+t^2}. \tag{12}$$

Así, aunque (10) define w como una función multivaluada de z, podemos representar z y w como funciones univaluadas o uniformes

de t. Las ecuaciones paramétricas (11) o (12) se dice que uniformizan la ecuación algebraica (10).

Para una ecuación $f(w,z) = 0$ de género 0, Clebsch [36] demostró que cada una de las variables puede expresarse como una función racional de un solo parámetro. Estas funciones racionales son funciones uniformizadoras. Cuando $f = 0$ se interpreta como una curva, entonces es llamada unicursal. Inversamente, si las variables w y z de $f = 0$ son expresables racionalmente en términos de un parámetro arbitrario, entonces $f = 0$ es de género 0.

En el mismo año,[37] Clebsch demostró que cuando $p = 1$, entonces w y z pueden expresarse como funciones racionales de los parámetros ξ y η donde η^2 es un polinomio de grado tres o cuatro en ξ. Entonces $f(w,z) = 0$, o la curva correspondiente es llamada bicursal, un término introducido por Cayley.[38] También es llamada elíptica, porque la ecuación $(dw/dz)^2 = \eta^2$ conduce a las integrales elípticas. También podemos decir que w y z se pueden expresar como funciones univaluadas doblemente periódicas de un solo parámetro α o como funciones racionales de $\wp(\alpha)$ donde $\wp(\alpha)$ es la función de Weierstrass. El resultado de Clebsch sobre la uniformización de curvas de género 1 por medio de funciones elípticas de un parámetro hizo posible establecer propiedades notables para tales curvas acerca de puntos de inflexión, cónicas osculatrices, tangentes desde un punto a una curva, y otros resultados, muchos de los cuales habían sido demostrado con anterioridad aunque con mayor dificultad.

Para las ecuaciones $f(w,z) = 0$ de género 2, Alexander von Brill (1842-1935) demostró [39] que las variables w y z se pueden expresar como funciones racionales de ξ y η donde η^2 es ahora un polinomio de quinto o sexto grado en ξ.

Por tanto, las funciones de género 0, 1 y 2 pueden uniformizarse. Para la función $f(w,z) = 0$ de género mayor que 2, la idea era emplear funciones más generales, a saber, funciones automorfas. En 1882, Klein [40] dio un teorema general de uniformización, pero la demostración no era completa. En 1883, Poincaré anunció [41] su teo-

[36] *Jour. für Math.*, 64, 1865, 43-65.
[37] *Jour. für Math.*, 64, 1865, 210-270.
[38] *Proc. Lon. Math. Soc.*, 4, 1871-1873, 347-352 = *Coll. Math. Papers*, 8, 181-187.
[39] *Jour. für March.*, 65, 1866, 269-283.
[40] *Math. Ann.*, 21, 1883, 141-218 = *Ges. math. Abh.*, 3, 630-710.
[41] *Bull. Soc. Math. de France*, 11, 1883, 112-125 = *Œuvres*, 4, 57-69.

rema general sobre uniformización pero él tampoco tenía una prueba completa. Tanto Klein como Poincaré continuaron trabajando duro para demostrar este teorema, pero no se obtuvo ningún resultado decisivo durante veinticinco años. En 1907 Poincaré [42] y Paul Koebe (1882-1945) [43] dieron independientemente una demostración de este teorema de uniformización. Koebe extendió después el resultado en muchas direcciones. Con el teorema sobre uniformización ya establecido rigurosamente ha llegado a ser posible un tratamiento perfeccionado de las funciones algebraicas y de sus integrales.

6. El enfoque de la geometría algebraica

Con la colaboración de Clebsch y Gordan empieza una nueva dirección de investigación en la geometría algebraica durante los años 1865-1870. Clebsch no estaba satisfecho con mostrar sólo la trascendencia de la obra de Riemann para las curvas. Buscaba ahora establecer la teoría de las integrales abelianas sobre la base de la teoría algebraica de las curvas. En 1865, él y Gordan reunieron sus esfuerzos en este trabajo y produjeron su *Theorie der Abelschen Funktionen (Teoría de las Funciones Abelianas,* 1866). Se debe apreciar que en esta época la teoría más rigurosa de Weierstrass sobre integrales abelianas no era conocida y el fundamento de Riemann —su prueba de existencia basada en el principio de Dirichlet— no solamente era extraño, sino que no estaba bien establecido. También en esta época hubo considerable entusiasmo por la teoría de los invariantes de formas algebraicas (o curvas) y por los métodos proyectivos como el primer estadio, por decirlo así, del tratamiento de las transformaciones birracionales.

Aunque el trabajo de Clebsch y Gordan fue una contribución a la geometría algebraica, no estableció una teoría puramente algebraica de la teoría de Riemann sobre las integrales abelianas. Sí que usaron métodos algebraicos y geométricos como opuestos a los métodos de la teoría de funciones de Riemann, pero también usaron resultados básicos de la teoría de funciones y los métodos de la teoría de funciones de Weierstrass. Además, tomaron algunos resultados acerca de las funciones racionales y el teorema del punto de

---

[42] *Acta Math.*, 31, 1908, 1-63 = *Œuvres*, 4, 70-139.

[43] *Math. Ann.*, 67, 1909, 145-224.

intersección como ciertos. Su contribución se sumó para que, partiendo de algunos resultados de teoría de funciones y usando métodos algebraicos, se obtuvieran nuevos resultados establecidos previamente por métodos de la teoría de funciones. Las transformaciones racionales fueron la esencia del método algebraico.

Ellos dieron la primera demostración *algebraica* de la invariancia del género p de una curva algebraica por las transformaciones racionales, usando como definición de p el grado y número de singularidades de $f = 0$. Entonces, usando el hecho de que p es el número de integrales linealmente independientes de primera clase sobre $f(x_1,x_2,x_3) = 0$ y que estas integrales son siempre finitas, mostraron que la transformación

$$\varrho x_i = \psi_i(y_1, y_2, y_3), \qquad i = 1, 2, 3,$$

transforma una integral de primera clase en una integral de primera clase de tal modo que p es invariante. También proporcionaron nuevas demostraciones del teorema de Abel (usando ideas y métodos de la teoría de funciones).

Su trabajo no era riguroso. En particular también ellos, en la tradición de Plücker, contaron las constantes arbitrarias para determinar el número de puntos de intersección de una C_m con una C_n. No investigaron las clases especiales de puntos dobles. La trascendencia de la obra de Clebsch-Gordan para la teoría de las funciones algebraicas consistió en expresar claramente en forma algebraica resultados tales como el teorema de Abel y usarlo en el estudio de las integrales abelianas. Colocaron la parte algebraica de la teoría de las integrales y funciones abelianas más en primer plano, y en particular establecieron la teoría de las transformaciones sobre fundamentos propios de ella.

Clebsch y Gordan habían planteado muchos problemas y dejado muchos huecos. Los problemas van en la dirección de nuevas investigaciones algebraicas de una teoría puramente algebraica de las funciones algebraicas. El trabajo con el enfoque algebraico fue continuado por Alexander von Brill y Max Noether a partir de 1871; su artículo principal fue publicado en 1874.[44] Brill y Noether basaron su teoría sobre un famoso teorema residual *(Restsatz)* que en sus

---

[44] «Über die algebraischen Funktionen und ihre Anwendung in der Geometrie». *Math. Ann.*, *7*, 1874, 269-310.

manos ocupó el lugar del teorema de Abel. También dieron una prueba algebraica del teorema de Riemann-Roch sobre el número de constantes que aparecen en las funciones algebraicas $F(w,z)$ que no se hacen infinitas en lugar alguno a excepción de m puntos predeterminados de una C_n. De acuerdo con este teorema, la función algebraica más general que satisface esta condición tiene la forma

$$F = C_1F_1 + C_2F_2 + \dots + C_\mu F_\mu + C_{\mu+1}$$

donde

$$\mu = m - p + \tau,$$

τ es el número de funciones ϕ linealmente independientes (de grado $n - 3$) que se anulan en los m puntos prescritos, y p es el género de la C_n. De esta manera, si la C_n es una C_4 sin puntos dobles, entonces $p = 3$ y las ϕ son líneas rectas. Para este caso cuando

$m = 1$, entonces $\tau = 2$ y $\mu = 1 - 3 + 2 = 0$;
$m = 2$, entonces $\tau = 1$, y $\mu = 2 - 3 + 1 = 0$;
$m = 3$, entonces $\tau = 1$ o 0. y $\mu = 1$ o 0.

Cuando $\mu = 0$ no hay ninguna función algebraica que se haga infinita en los puntos dados. Cuando $m = 3$, existe una y sólo una función, siempre que los tres puntos dados estén sobre una línea recta. Si los tres puntos caen en una recta $v = 0$, esta recta corta a la C_4 en un cuarto punto. Escogemos una recta $u = 0$ que pase por este punto y entonces $F_1 = u/v$.

Este trabajo reemplaza la determinación de Riemann de la función algebraica más general que tenga puntos dados en los cuales se haga infinita. También el resultado de Brill-Noether trasciende el punto de vista proyectivo en la medida en que trata de la geometría de los puntos sobre la curva C_n dada por $f = 0$, cuyas relaciones mutuas no son alteradas por una transformación birracional inyectiva. Así, por primera vez, se establecieron los teoremas sobre puntos de intersección de curvas de manera algebraica. Se prescindió de contar las constantes como método.

El trabajo en geometría algebraica continuó con la investigación

detallada de curvas espaciales algebraicas por Noether [45] y Halphen.[46] Cualquier curva espacial C puede ser proyectada birracionalmente en una curva plana C_1. Todas las C_1 que se obtienen a partir de C tienen el mismo género. Por tanto, el género de C se define como el de cualquiera de dichas C_1, y el género de C es invariante bajo una transformación birracional del espacio.

El tema que ha recibido la mayor atención durante el transcurso del tiempo es el estudio de las singularidades de las curvas algebraicas planas. Hasta 1871, la teoría de las funciones algebraicas considerada desde el punto de vista algebraico se había limitado al estudio de curvas que tengan puntos dobles distintos o separados, y en el peor de los casos solamente cúspides *(Ruckkehrpunkte)*. Se creía que las curvas con singularidades más complicadas podían ser tratadas como casos límites de curvas con puntos dobles. Pero el procedimiento efectivo de paso al límite era vago y carecía de rigor y unidad. La culminación del trabajo sobre singularidades son dos famosos teoremas sobre transformaciones. El primero afirma que toda curva algebraica plana irreducible puede ser transformada por medio de una transformación de Cremona en una que no tenga más puntos singulares que puntos múltiples con tangentes distintas. El segundo afirma que, por medio de una transformación birracional solamente en la curva, toda curva algebraica plana irreducible puede transformarse en otra que tenga solamente puntos dobles con tangentes distintas. La reducción de las curvas a estas formas más sencillas facilita la aplicación de muchos de los métodos de la geometría algebraica.

Sin embargo, las numerosas pruebas de estos teoremas, especialmente del segundo, han sido incompletas o al menos criticadas por los matemáticos (excepto por el autor). Realmente, hay dos casos del segundo teorema, curvas reales en el plano proyectivo y curvas en el sentido de la teoría de las funciones complejas, donde x y y se toman cada una de ellas en un plano complejo. Noether [47] usó en 1871 una sucesión de transformaciones cuadráticas que son uno-a-uno en todo el plano para demostrar el primer teorema. Generalmente se le atribuye la prueba, pero realmente él sólo indicó una demostración que fue perfeccionada y modificada por muchos au-

---

[45] *Jour. für Math.*, 93, 1882, 271-318.

[46] *Jour. de l'Ecole Poly.*, Cahier 52, 1882, 1-200 = *Œuvres*, 3, 261-455.

[47] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1871, 267-278.

tores.[48] Kronecker, usando el análisis y el álgebra, desarrolló un método para probar el segundo teorema. Comunicó este resultado verbalmente a Riemann y Weierstrass en 1858, dio conferencias sobre él a partir de 1870 y lo publicó en 1881.[49] El método usaba transformaciones birracionales, las cuales, con la ayuda de la ecuación de la curva plana dada, son uno-a-uno, y transformaban el caso singular en el «regular»; esto es, los puntos singulares se convierten precisamente en puntos dobles con tangentes distintas. El resultado, sin embargo, no fue enunciado por Kronecker y sólo se encuentra implícito en su obra.

Este segundo teorema, con la intención de que todos los puntos múltiples se puedan reducir a puntos dobles por medio de transformaciones birracionales sobre las curvas, fue enunciado primero explícitamente y demostrado por Halphen en 1884.[50] Se han dado muchas otras pruebas, pero ninguna es universalmente aceptada.

7. El enfoque aritmético

Además del enfoque transcendente y del de la geometría algebraica, está el llamado enfoque aritmético de las curvas algebraicas, el cual es, sin embargo, al menos en concepto, puramente algebraico. Este enfoque es realmente un grupo de teorías que difieren grandemente en detalle, pero que tienen en común la construcción y el análisis de los integrandos de las tres clases de integrales abelianas. Este enfoque fue desarrollado por Kronecker en sus conferencias,[51] por Weierstrass en las suyas de 1875-1876 y por Dedekind y Heinrich Weber en un artículo escrito conjuntamente.[52] El enfoque se presenta en su totalidad en el texto de Kurt Hensel y Georg Landsberg: *Theorie der algebraischen Funktionen einer Variabeln (Teoría de las Funciones Algebraicas de una Variable,* 1902).

---

[48] Véase también Noether, *Math. Ann.*, 9, 1876, 166-182; y 23, 1884, 311-358.

[49] *Jour. für Math.*, 91, 1881, 301-334 = *Werke*, 2, 193-236.

[50] Reproducido en un apéndice a una edición francesa (1884) del *Higher Plane Curves* de G. Salmon y en el *Traité d'analyse* de E. Picard, 2, 1893, 364 y sgs. = *Œuvres*, 4, 1-93.

[51] *Jour. für Math.*, 91, 1881, 301-334 y 92, 1882, 1-122 = *Werke*, 2, 193-387.

[52] «Theorie der algebraischen Funktionen einer Veränderlichen». *Jour. für Math.*, 92, 1882, 181-290 = *Werke* de Dedekind, I, 238-350.

La idea central de este enfoque proviene del trabajo sobre los números algebraicos de Kronecker y Dedekind y utiliza una analogía entre los enteros algebraicos de un cuerpo de números algebraicos y las funciones algebraicas en la superficie de Riemann de una función compleja. En la teoría de los números algebraicos se empieza con una ecuación polinomial irreducible $f(x) = 0$ con coeficientes enteros. Lo análogo para la geometría algebraica es una ecuación polinomial irreducible $f(\zeta,z) = 0$, cuyos coeficientes en las potencias de ζ son polinomios en z (digamos con coeficientes reales). En la teoría de números se considera entonces el cuerpo $R(x)$ generado por los coeficientes de $f(x) = 0$ y una de sus raíces. En geometría se considera el cuerpo de todas las $R(\zeta,z)$ que són algebraicas y univaluadas en la superficie de Riemann. Se considera entonces en la teoría de números los números algebraicos enteros. A éstos les corresponden las funciones algebraicas $G(\zeta,z)$ que son enteras, esto es, se hacen infinitas solamente en $z = \infty$. La descomposición de los enteros algebraicos en factores primos reales y unidades corresponde respectivamente a la descomposición de la $G(\zeta,z)$ en factores tales que cada uno de ellos se anula solamente en un punto de la superficie de Riemann y en factores que no se anulan en ningún punto, respectivamente. Donde Dedekind introdujo ideales en la teoría de números para discutir la divisibilidad, en el análogo geométrico se reemplaza un factor de una $G(\zeta,z)$ que se anula en un punto de la superficie de Riemann por la colección de todas las funciones del cuerpo de $R(\zeta,z)$ que se anulan en ese punto. Dedekind y Weber usaron este método aritmético para tratar el cuerpo de las funciones algebraicas y obtuvieron los resultados clásicos.

Hilbert [53] continuó lo que es esencialmente el enfoque algebraico o aritmético de Dedekind y Kronecker a la geometría algebraica. Un teorema principal, el *Nullstellensatz* de Hilbert, dice que toda estructura algebraica (figura) de extensión arbitraria en un espacio de un número arbitrario de variables homogéneas $x_1, \dots x_n$ siempre puede representarse por un número finito de ecuaciones homogéneas

$$F_1 = 0, F_2 = 0, \dots, F_\mu = 0,$$

[53] «Über die Theorie der algebraischen Formen». *Math. Ann.*, 36, 1890, 473-534 = *Ges. Abh.*, 2, 199-257.

de modo que la ecuación de cualquier otra estructura que contenga a la original puede representarse por

$$M_1F_1 + \ldots + M_\mu E_\mu = 0,$$

donde las M son formas enteras homogéneas arbitrarias cuyo grado debe escogerse de modo tal que el lado izquierdo de la ecuación sea él mismo homogéneo.

Siguiendo a Dedekind, Hilbert llamó a la colección de los M_iF_i un módulo (el término que ahora se emplea es ideal y módulo es algo más general). Se puede enunciar el resultado de Hilbert así: toda estructura algebraica de R_n determina la anulación de un módulo finito.

8. La geometría algebraica de superficies

Casi desde el inicio de la investigación en la geometría algebraica de las curvas, se investigó también la teoría de las superficies. Aquí también el trabajo se dirigió hacia el estudio de los invariantes bajo transformaciones lineales y birracionales. Igual que la ecuación $f(x,y) = 0$, la ecuación polinomial $f(x,y,z) = 0$ tiene una doble interpretación. Si x, y y z toman valores reales, entonces la ecuación representa una superficie bidimensional en el espacio tridimensional. Sin embargo, si estas variables toman valores complejos, entonces la ecuación representa una variedad tetradimensional en un espacio de seis dimensiones.

El enfoque de la geometría algebraica de las superficies fue análogo al de las curvas. Clebsch empleó métodos de la teoría de funciones e introdujo [54] integrales dobles que hacen el papel de las integrales abelianas en la teoría de las curvas. Clebsch observó que para una superficie algebraica $f(x,y,z) = 0$ de grado m con puntos múltiples aislados y rectas múltiples ordinarias, ciertas superficies de grado $m - 4$ tendrían que jugar el papel que las curvas adjuntas de grado $m - 3$ juegan con respecto a una curva de grado m. Dada una función racional $R(x,y,z)$, donde x, y y z están relacionadas por $f(x,y,z) = 0$, si se investigan las integrales dobles

[54] *Comp. Rend.*, *67*, 1868, 1238-1239.

$$\int\int R(x, y, z)dxdy,$$

que permanecen siempre finitas cuando las integrales se extienden sobre una región bidimensional de la superficie tetradimensional, se encuentra que son de la forma

$$\int\int \frac{Q(x, y, z)}{f_z}\, dx\, dy$$

donde Q es un polinomio de grado $m - 4$. $Q = 0$ es una superficie adjunta que pasa por las rectas múltiples de $f = 0$ y tiene una recta múltiple de orden $k - 1$ al menos en cada recta múltiple de f *de* orden k y tiene un punto múltiple de orden $q - 2$ al menos en cada punto múltiple aislado de orden q. Tal integral es llamada una integral doble de primera clase. El número de integrales linealmente independientes de esta clase, que es el número de constantes esenciales en $Q(x,y,z)$, es llamado el género geométrico p_g de $f = 0$. Si la superficie no tiene rectas múltiples de puntos

$$p_g = \frac{(m-1)(m-2)(m-3)}{6}.$$

Max Noether[55] y Hieronymus G. Zeuthen (1839-1920)[56] demostraron que p_g es invariante por las transformaciones birracionales de la superficie (no de todo el espacio).

Hasta este punto la analogía con la teoría de las curvas es buena. Las integrales dobles de primera clase son análogas a las integrales abelianas de primera clase. Pero se manifiesta ahora una primera diferencia. Es necesario calcular el número de constantes esenciales en los polinomios Q de grado $m - 4$ que en puntos múltiples de la superficie se comportan de manera tal que la integral permanece finita. Pero por medio de una fórmula se puede encontrar el número preciso de condiciones de este modo involucradas solamente para un polinomio de grado N suficientemente grande. Si en esta fórmula se pone $N = m - 4$ se podría encontrar un número diferente de p_g.

[55] *Math. Ann.*, 2, 1870, 293-316.
[56] *Math. Ann.*, 4, 1871, 21-49.

Cayley [57] llamó a este nuevo número el género numérico (aritmético) p_n de la superficie. El caso más general es cuando $p_n = p_g$. Cuando no se cumple la igualdad se tiene $p_n < p_g$ y la superficie es llamada irregular, si no es llamada regular. Después Zeuthen [58] y Noether [59] establecieron la invariancia del número p_n cuando no es igual a p_g.

Picard [60] desarrolló una teoría de integrales dobles de segunda clase. Estas son las integrales que se hacen infinitas a la manera de

$$\int\int \left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right) dx\, dy \tag{13}$$

donde U y V son funciones racionales de x, y y z y $f(x,y,z) = 0$. El número de integrales diferentes de segunda clase, diferentes en el sentido de que ninguna combinación lineal de estas integrales se reduce a la forma (13), es finito; este es un invariante birracional de la superficie $f = 0$. Pero no es cierto aquí, como en el caso de las curvas, que el número de integrales abelianas distintas de segunda clase es $2p$. Este nuevo invariante de las superficies algebraicas no parece estar vinculado con el género numérico o el geométrico.

Hasta ahora se ha logrado mucho menos para la teoría de las superficies que para las curvas. Una razón es que las posibles singularidades de las superficies son mucho más complicadas. Se tiene el teorema de Picard y Georges Simart demostrado por Beppo Levi (1875-1928) [61] de que cualquier superficie algebraica (real) puede ser transformada birracionalmente en una superficie sin singularidades, la cual, sin embargo, debe estar en un espacio de cinco dimensiones. Pero este teorema no resulta ser demasiado útil.

En el caso de las curvas, el número invariante, el género p, es susceptible de ser definido en términos de las características de la curva o la conectividad de la superficie de Riemann. En el caso de $f(x,y,z) = 0$, el número que caracteriza los invariantes birracionales aritméticos es desconocido [62]. No intentaremos describir más a fon-

---

[57] *Phil. Trans.*, 159, 1869, 201-229 = *Coll. Math. Papers*, 6, 329-358; y *Math. Ann.*, 3, 1871, 526-529 = *Coll. Math. Papers*, 8, 394-397.

[58] *Math. Ann.*, 4, 1871, 21-49.

[59] *Math. Ann.*, 8, 1875, 495-533.

[60] *Jour. de Math.* (5), 5, 1899, 5-54, y artículos posteriores.

[61] *Annali di Mat.* (2), 26, 1897, 219-253.

[62] Las variedades involucradas no pueden ser caracterizadas, ni siquiera topológicamente.

do los pocos resultados limitados para la geometría algebraica de las superficies.

El tema de estudio de la geometría algebraica abarca ahora el estudio de figuras de mayores dimensiones («manifolds» o variedades) definidas por una o más ecuaciones algebraicas. Más allá de la generalización en esta dirección, otra extensión, a saber, el uso de coeficientes más generales en las ecuaciones que las definen, también ha sido llevada a cabo. Estos coeficientes pueden ser elementos de un anillo o cuerpo abstractos y se aplican los métodos del álgebra abstracta. Los diversos métodos de cultivar la geometría algebraica, al igual que la formulación algebraica abstracta introducida en el siglo XX, han llevado a agudas diferencias en el lenguaje y métodos de enfoque, de modo que una clase de investigadores encuentra muy difícil entender a otras. En este siglo, el interés ha sido puesto en el enfoque algebraico abstracto. Parece que sí ofrece formulaciones y teoremas precisos, y de ese modo las demostraciones zanjan muchas controversias acerca del significado y corrección de los resultados anteriores. Sin embargo, parece que mucho del trabajo hecho tiene mucha más relación con el álgebra que con la geometría.

Bibliografía

Barker, H. F.: «On Some Recent Advances in the Theory of Algebraic Surfaces». *Proc. Lon. Math. Soc.* (2), 12, 1912-1913, 1-40.

Berzolari, L.: «Allgemeine Theorie der höheren ebenen algebraischen Kurven». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III C4, 313-455.

— «Algebraische Transformationen und Korrespondenzen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III, 2, 2.ª parte B, 1781-2218. Util por los resultados sobre figuras de mayores dimensiones.

Bliss, G. A.: «The Reduction of Singularities of Plane Curves by Birational Transformations». *Amer. Math. Soc. Bull.*, 29, 1923, 161-183.

Brill, A., y Noether, M.: «Die Entwicklung der Theorie der algebraischen Funktionen». *Jahres. der Deut. Math.-Verein.*, 3, 1892-1893, 107-565.

Castelnuovo, G., y Enriques, F.: «Die algebraischen Flächen vom Gesichtspunkte der birationalen Transformationen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III C6b, 674-768.

— «Sur quelques récents résultats dans la théorie des surfaces algébriques». *Math. Ann.*, 48, 1897, 241-316.

Cayley, A.: *Collected Mathematical Papers.* Johnson Reprint. Corp., 1963, vols. 2, 4, 6, 7, 10, 1891-1896.

Clebsch, R. F. A.: «Versuch einer Darlegung und Würdigung seiner Wissenschaftlichen Leistungen». *Math. Ann.*, 7, 1874, 1-55. Un artículo escrito por amigos de Clebsch.

Coolidge, Julian L.: *A History of Geometrical Methods*, Dover (reimp.), 1963, 195-230, 278-292.

Cremona, Luigi: *Opere mathematiche*, 3 vols., Ulrico Hoepli, 1914-1917.

Hensel, Kurt, y Georg Landsberg: *Theorie der algebraischen Funktionen einer Variabeln* (1902), Chelsea (reimp.), 1965, 694-702 en particular.

Hilbert, David: *Gesammelte Abhandlungen*. Julius Springer, 1933, vol. 3.

Klein, Felix: *Vorlesungen über die Entwicklung der Mathematik im 19 Jahrhundert*, 1, 155-166, 295-319; 2, 2-26; Chelsea (reimp.), 1950.

Meyer, Franz W.: «Berich über den gegenwärtigen Stand der Invariantentheorie». *Jahres. der Deut. Math.-Verein.*, 1, 1890-1991, 79-292.

National Research Council: *Selected Topics in Algebraic Geometry*, Chelsea (reimp.), 1970.

Noether, Emmy: «Die arithmetische Theorie der algebraischen Funktionen einer Veränderlichen in ihrer Beziehung zu den übrigen Theorien and zu der Zahlentheorie». *Jahres. der Deut. Math.-Verein.*, 28, 1919, 182-203.

Capítulo 40
LA INTRODUCCION DEL RIGOR EN EL ANALISIS

> Pero sería un serio error pensar que sólo se puede encontrar certeza en las demostraciones geométricas o en el testimonio de los sentidos.
>
> A. L. CAUCHY

1. Introducción

Hacia el año 1800, los matemáticos empezaron a interesarse por la imprecisión de los conceptos y demostraciones de vastas ramas del análisis. El propio concepto de función no estaba claro; el uso de las series sin referencia a la convergencia y divergencia había producido paradojas y desacuerdos; la controversia sobre las representaciones de las funciones por medio de series trigonométricas había introducido mayor confusión; y, desde luego, las nociones fundamentales de derivada e integral nunca habían sido definidas correctamente. Todas estas dificultades causaron finalmente insatisfacción con el *status* lógico del análisis.

Abel, en una carta de 1826 al profesor Christoffer Hansteen,[1] se quejaba de «la tremenda oscuridad que incuestionablemente encuentra uno en el análisis. Carece completamente de todo plan y sistema y el hecho de que tantos lo hayan podido estudiar es notable. Lo peor del caso es que nunca ha sido tratado rigurosamente. Hay muy pocos teoremas en el análisis superior que se hayan demostrado de

[1] *Œuvres*, 2, 263-265.

una manera lógicamente sostenible. En todas partes encuentra uno esta manera miserable de concluir lo general partiendo de lo especial y es extremadamente peculiar que tal procedimiento haya llevado a tan pocas de las así llamadas paradojas».

Varios matemáticos se resolvieron a poner orden en todo este caos. Las cabezas de lo que frecuentemente es llamado el movimiento crítico decidieron reconstruir el análisis solamente sobre la base de los conceptos aritméticos. Los principios del movimiento coinciden con la creación de la geometría no euclídea. Un grupo totalmente diferente, excepto Gauss, se involucró en esta última actividad y, por tanto, es difícil rastrear alguna relación directa entre ella y la decisión de fundar el análisis en la aritmética. Tal vez se llegó a esa decisión porque la esperanza de fundar el análisis sobre la geometría, lo cual muchos matemáticos del siglo XVII afirmaron con frecuencia que sí se podía hacer, fue desechado por la complejidad creciente de los desarrollos en análisis durante el siglo XVIII. Sin embargo, Gauss ya había expresado sus dudas en cuanto a la verdad de la geometría euclídea en 1799 y en 1817 decidió que la verdad residía únicamente en la aritmética. Más aún, incluso durante la obra inicial de Gauss y otros en la geometría no euclídea, ya se habían notado imperfecciones en el desarrollo de Euclides. Por tanto, es muy probable que ambos factores causaran desconfianza en la geometría y apresuraran la decisión de fundar el análisis sobre conceptos aritméticos. Esto fue ciertamente lo que las cabezas del movimiento crítico se comprometieron a llevar a cabo.

El análisis riguroso empieza con la obra de Bolzano, Cauchy, Abel y Dirichlet, y Weierstrass lo fomentó. Cauchy y Weierstrass son los más conocidos en esta relación. Las obras básicas de Cauchy sobre los fundamentos del análisis son: *Cours d'analyse algébrique* (Curso de Análisis Algebraico)[2], *Résumé des leçons sur le calcul infinitesimal* (Resumen de las Lecciones sobre el Cálculo Infinitesimal)[3] y *Leçons sur le calcul différentiel* (Lecciones sobre el Cálculo Diferencial).[4] Realmente, el rigor de Cauchy en estas obras es impreciso de acuerdo con los criterios modernos. Usó frases tales como «se acerca indefinidamente», «tan poco como se desee», «últimas razones de incrementos infinitamente pequeños» y «una variable se

[2] 1821, *Œuvres* (2), III.
[3] 1823, *Œuvres* (2), IV, 1-261.
[4] 1829, *Œuvres* (2), IV, 265-572.

acerca a su límite». Sin embargo, si uno compara la *Théorie des fonctions analytiques* (Teoría de las Funciones Analíticas)[6] de Lagrange y sus *Leçons sur le calcul des fonctions* (Lecciones sobre el Cálculo de las Funciones)[6] y el libro de Lacroix, que tanta influencia tuvo, *Traité du calcul différentiel et du calcul intégral* (Tratado de Cálculo Diferencial y de Cálculo Integral)[7] junto con el *Cours d'analyse algébrique* de Cauchy, uno empieza a ver la sorprendente diferencia entre las matemáticas del siglo XVIII y las del XIX. Lagrange, en particular, era puramente formal. Operaba con expresiones simbólicas. Los conceptos subyacentes de límite, continuidad, etc. no están ahí.

Cauchy es muy explícito en su introducción a la obra de 1821 en cuanto a que busca dar rigor al análisis. Señala que el uso libre para todas las funciones de las propiedades que se cumplen para las funciones algebraicas y el uso de series divergentes no están justificados. Aunque la obra de Cauchy no fue sino un paso en la dirección del rigor, él mismo creía y establece en su *Résumé* que había llegado a lo último en rigor dentro del análisis. Sí proporcionó los comienzos de demostraciones precisas de teoremas y limitó las afirmaciones adecuadamente, al menos para las funciones elementales. Abel, en su artículo de 1826 sobre serie binomiales (sec. 5) alabó este logro de Cauchy: «La distinguida obra [el *Cours d'analyse*] debiera ser leída por todo aquel que ame el rigor dentro de las investigaciones matemáticas.» Cauchy abandonó las representaciones explícitas de Euler y las series de potencias de Lagrange e introdujo nuevos conceptos para tratar las funciones.

2. Las funciones y sus propiedades

Los matemáticos del siglo XVIII creían en su totalidad que una función debía tener la misma expresión analítica en todas partes. Durante la parte final del siglo, en gran medida como consecuencia de la controversia sobre el problema de la cuerda vibrante, Euler y Lagrange permitieron las funciones que tienen diferentes expresiones en diferentes dominios y usaron la palabra continua donde se man-

---

[5] 1797; 2.ª ed. 1813 = *Œuvres*, 9.

[6] 1801; 2.ª ed. 1806 = *Œuvres*, 10.

[7] 3 vols., 1.ª ed., 1797-1800; 2.ª ed., 1810-1819.

tenía la misma expresión y discontinua en puntos donde la expresión cambiaba de forma (aunque en el sentido moderno toda la función podía ser continua). Mientras Euler, d'Alembert y Lagrange tenían que reconsiderar el concepto de función, no llegaron a ninguna definición ampliamente aceptada y tampoco resolvieron el problema de qué funciones podían representarse por medio de series trigonométricas. Sin embargo, la expansión gradual en la variedad y uso de las funciones forzó a los matemáticos a aceptar un concepto más amplio.

Gauss, en su trabajo inicial, interpretaba por función una expresión cerrada (analítica finita) y cuando hablaba de la serie hipergeométrica $F(\alpha,\beta,\gamma,x)$ como una función de α,β,γ y x la explicaba señalando, «en tanto se la pueda tratar como función». Lagrange ya había usado un concepto más amplio al tratar las series de potencias como funciones. En la segunda edición de su *Mécanique analytique* (1811-1815) usó la palabra función para casi cualquier clase de dependencia de una o más variables. Aún Lacroix, en su *Traité* de 1797, había introducido ya una noción más amplia. En la introducción dice: «Toda cantidad cuyo valor depende de una o varias otras es llamada una función de estas últimas, ya sea que uno conozca o no por medio de qué operaciones es necesario pasar de las últimas a la primera cantidad.» Lacroix da como un ejemplo de raíz de una ecuación de quinto grado como una función de sus coeficientes.

La obra de Fourier abrió aún más ampliamente la discusión sobre lo que es una función. Por un lado, insistió en que las funciones no necesitan ser representables por una expresión analítica. En su *Teoría Analítica del Calor* [8] dice: «En general la función $f(x)$ representa una sucesión de valores u ordenadas, cada una de las cuales es arbitraria. ... No suponemos que estas ordenadas estén sujetas a una ley común; se suceden una a la otra de cualquier manera, sea la que fuere...» De hecho, sólo trató funciones con un número finito de discontinuidades en cualquier intervalo finito. Por otro lado, hasta cierto grado, Fourier estaba apoyando la aseveración que una función debe ser representable por una expresión analítica, aunque esta expresión fuese una serie de Fourier. En cualquier caso, la obra de Fourier perturbó la creencia del siglo XVIII de que todas las funciones, en el peor de los casos, eran extensiones de funciones algebraicas. Las funciones algebraicas, y aun las funciones trascendentes ele-

[8] Traducción en inglés, p. 430, Dover (reimp.), 1955.

mentales, ya no fueron el prototipo de funciones. Como las propiedades de las funciones algebraicas ya no podían cumplirse para todas las funciones, surgió entonces la cuestión en cuanto a lo que realmente se quiere decir por función, por continuidad, diferenciabilidad, integrabilidad y otras propiedades.

En la reconstrucción positiva del análisis, que muchos matemáticos se propusieron llevar a cabo, el sistema de los números reales se dio por sentado. No se hizo ningún intento por analizar esta estructura o por construirla lógicamente. Aparentemente, los matemáticos sentían que trabajaban sobre bases seguras en cuanto a lo concerniente a este área.

Cauchy empieza su trabajo de 1821 con la definición de una variable. «Se llama variable a una cantidad que se considera tiene que tomar sucesivamente muchos valores diferentes unos de los otros.» En cuanto al concepto de función: «Cuando se relacionan cantidades variables entre ellas de modo que estando dado el valor de una de éstas, se puedan determinar los valores de todas las otras, ordinariamente se concibe a estas cantidades diversas expresadas por medio de la que está entre ellas, la cual entonces toma el nombre de variable independiente; y las otras cantidades expresadas por medio de la variable independiente son aquellas que uno llama funciones de esta variable.» Cauchy también es explícito en que una serie infinita es una manera de especificar una función. Sin embargo, no se requiere una expresión analítica para una función.

En un artículo sobre series de Fourier, al cual retornaremos posteriormente, «Uber die Darstellung ganz willkürlicher Functionen durch Sinus-und Cosinusreihen» («Sobre la Representación de Funciones Completamente Arbitrarias por Series de Senos y Cosenos»),[9] Dirichlet dio la definición de una función (univaluada) que es la que más frecuentemente se emplea ahora, a saber, que y es una función de x cuando a cada valor de x en un intervalo dado le corresponde un único valor de y. Agregó que no importa si en todo este intervalo y depende de x de acuerdo a una ley o más, o si la dependencia de y de x puede expresarse por medio de operaciones matemáticas. De hecho, en 1829[10] dio el ejemplo de una función de x que tiene el valor c para todos los valores racionales de x y el valor d para todos los valores irracionales de x.

---

[9] *Repertorium der Physik*, 1, 1837, 152-174 = *Werke*, 1, 135-160.
[10] *Jour. für Math.*, 4, 1829, 157-169 = *Werke*, 1, 117-132.

Hankel señala que los mejores libros de texto de al menos la primera mitad del siglo no sabían qué hacer con el concepto de función. Algunos definían una función esencialmente en el sentido de Euler; otros requerían que y variara con x de acuerdo con alguna ley, pero no explicaban lo que quería decir ley; algunos usaban la definición de Dirichlet; y aún otros no dieron definición alguna. Pero todos dedujeron consecuencias a partir de sus definiciones que no estaban implicadas lógicamente por las definiciones.

La distinción adecuada entre continuidad y discontinuidad surgió gradualmente. El estudio cuidadoso de las propiedades de las funciones fue iniciado por Bernhard Bolzano (1781-1848), un sacerdote, filósofo y matemático de Bohemia. Bolzano fue llevado a este trabajo cuando trataba de proporcionar una demostración puramente aritmética del teorema fundamental del álgebra en lugar de la primera demostración de Gauss (1799), la cual usaba ideas geométricas. Bolzano tenía los conceptos correctos para el establecimiento del cálculo (a excepción de una teoría de los números reales), pero su trabajo pasó desapercibido durante medio siglo. Negó la existencia de números infinitamente pequeños (infinitésimos) y de números infinitamente grandes, ambos habían sido utilizados por los autores del siglo XVIII. En un libro de 1817, cuyo largo título empieza con *Rein analytischer Beweis* (véase la bibliografía), Bolzano dio la definición apropiada de continuidad, a saber, $f(x)$ es continua en un intervalo si en cualquier x del intervalo la diferencia $f(x + w) - f(x)$ se puede hacer tan pequeña como se desee tomando w suficientemente pequeña. Prueba que los polinomios son continuos.

Cauchy, también, abordó las nociones de límite y continuidad. Como con Bolzano, el concepto de límite estaba basado en consideraciones puramente aritméticas. En el *Cours* (1821), Cauchy dice: «Cuando los valores sucesivos asignados a una variable se acercan indefinidamente a un valor fijo de modo que terminan por diferir de él por tan poco como se desee, este último es llamado el límite de los otros. Así, por ejemplo, un número irracional es el límite de diversas fracciones que toman valores cada vez más aproximados a él.» Este ejemplo fue un poco desafortunado porque muchos lo tomaron como una definición de los números irracionales en términos de límite, siendo así que el límite podría no tener significado si los irracionales no estuvieran ya presentes. Cauchy lo omitió en sus obras de 1823 y 1829.

En el prefacio a su trabajo de 1821, Cauchy dice que para hablar

de la continuidad de las funciones debe dar a conocer las propiedades principales de las cantidades infinitamente pequeñas. «Se dice [*Cours*, pág. 5] que una cantidad variable se hace infinitamente pequeña cuando su valor numérico decrece indefinidamente de modo tal que converge al límite 0.» Llama a tales variables infinitésimos. Así clarifica Cauchy la noción de infinitésimo de Leibniz y la libera de ataduras metafísicas. Cauchy continúa: «Se dice que una cantidad variable se hace infinitamente grande cuando su valor numérico se incrementa indefinidamente de manera tal que converge al límite ∞.» Sin embargo, ∞ no significa una cantidad fija, sino algo indefinidamente grande.

Ahora, Cauchy está preparado para definir la continuidad de una función. En el *Cours* (págs. 34-35) dice: «Sea $f(x)$ una función de la variable x, y supóngase que, para cada valor de x que se encuentre entre dos límites [cotas] dados, esta función toma constantemente un valor finito y único. Si, empezando con un valor de x contenido entre estos límites, se asigna a la variable x un incremento infinitamente pequeño, la función misma tomará como incremento la diferencia $f(x + \alpha) - f(x)$, el cual dependerá al mismo tiempo de la nueva variable α y del valor de x. Garantizado esto, la función $f(x)$ será, entre los dos límites asignados a la variable x, una función *continua* de la variable si, para cada valor de x que se encuentre entre estos dos límites, el valor numérico de la diferencia $f(x + \alpha) - f(x)$ decrece indefinidamente con el de α. En otras palabras, *la función $f(x)$ permanecerá continua con respecto a x entre los dos límites dados, si, entre estos límites, un incremento infinitamente pequeño de la variable produce siempre un incremento infinitamente pequeño de la función misma.*

«También decimos que la función $f(x)$ es una función continua de x en la vecindad de un valor particular asignado a la variable x, siempre que sea continua [la función] entre esos dos límites de x, no importe cuán cercanos estén, los cuales encierren el valor en cuestión.» Entonces dice que $f(x)$ es discontinua en x si no es continua en todo el intervalo alrededor de x.

En el *Cours* (pág. 37) Cauchy afirmó que si una función de varias variables es continua en cada una por separado, entonces es una función continua de todas las variables. Esto no es correcto.

Durante todo el siglo XIX se exploró la noción de continuidad y los matemáticos aprendieron más acerca de ella, a veces produciendo resultados que los asombraban. Darboux dio un ejemplo de una

función que tomaba todos los valores intermedios entre dos valores dados al pasar de $x = a$ a $x = b$, pero que no era continua. Así, una propiedad básica de las funciones continuas no es suficiente para asegurar la continuidad.[11]

El trabajo de Weierstrass sobre la rigorización del análisis mejoró el de Bolzano, Abel y Cauchy. También buscó la manera de evitar la intuición y de basarse en conceptos aritméticos. Aunque realizó este trabajo durante los años 1841-1856, cuando era maestro de escuela de educación media, mucho de ello no se conoció hasta 1859 cuando empezó a sentar cátedra en la Universidad de Berlín.

Weierstrass atacó la frase «una variable se acerca a un límite», la cual, desafortunadamente, sugiere tiempo y movimiento. Interpreta una variable simplemente como una letra que se usa para denotar cualquier valor de los de un conjunto dado que se le puede asignar a la letra. Así se elimina la idea de movimiento. Una variable continua es tal que si x es cualquier valor del conjunto de valores de la variable y δ cualquier número positivo, hay otros valores de la variable en el intervalo $(x_0 - \delta, x_0 + \delta)$.

Para eliminar la vaguedad en la frase «se hace y permanece menor que cualquier cantidad dada», que Bolzano y Cauchy usaron en sus definiciones de continuidad y límite de una función, Weierstrass dio la definición ahora aceptada de que $f(x)$ es continua en $x = x_0$ si dado cualquier número positivo ε, existea una δ tal que para toda x en el intervalo $|x - x_0|<\delta$, $|f(x) - f(x_0)|<\varepsilon$. Una función $f(x)$ tiene un límite L en $x = x_0$ si se cumple la misma afirmación, pero con L reemplazando a $f(x_0)$. Una función $f(x)$ es continua en un intervalo de valores x si es continua en cada x del intervalo.

Durante los años en que la noción de continuidad misma se estaba refinando, los esfuerzos por establecer el análisis rigurosamente exigían la demostración de muchos teoremas sobre funciones continuas que habían sido aceptados intuitivamente. Bolzano, en su publicación de 1817, buscó demostrar que si $f(x)$ es negativa para $x = a$ y positiva para $x = b$, entonces $f(x)$ tiene un cero entre a y b. Consideró la sucesión de funciones (para x fija)

$$F_1(x),\ F_2(x),\ F_3(x),\ \ldots,\ F_n(x),\ \ldots \qquad (1)$$

---

[11] Considérese $y = \text{sen}\,(1/x)$ para $x \neq 0$ e $y = 0$ para $x = 0$. Esta función va de todos los valores desde uno que tome para un valor negativo de x hasta uno que tome para un valor positivo de x. Sin embargo, no es continua en $x = 0$.

e introdujo el teorema que afirma que si para n suficientemente grande podemos hacer la diferencia $F_{n+r} - F_n$ menor que cualquier cantidad dada, por grande que sea r, entonces existe una magnitud fija X tal que la sucesión se acerca cada vez más a X, y ciertamente tanto como se desee. Su determinación de la cantidad X fue algo oscura, porque no tenía una teoría particular clara del sistema de los números reales y de los números irracionales sobre la cual basarse. Sin embargo, tuvo la idea de lo que ahora llamamos la condición de Cauchy para la convergencia de una sucesión (ver más adelante).

En el desarrollo de la demostración, Bolzano estableció la existencia de una cota superior mínima para un conjunto acotado de números reales. Su enunciado preciso es: si una propiedad M no se aplica a todos los valores de una cantidad variable x, sino a todos aquellos que son más pequeños que una cierta u, siempre hay una cantidad U que es la mayor de las que se puede afirmar que toda x más pequeña posee la propiedad M. La esencia de la demostración de Bolzano de este lema consistió en dividir el intervalo acotado en dos partes y seleccionar una parte particular que contenga un número infinito de miembros del conjunto. En seguida repite el proceso hasta que finalmente encierra el número que es la mínima cota superior del conjunto dado de números reales. Este método fue usado por Weierstrass en los sesenta, con el debido tributo a Bolzano, para demostrar lo que ahora se llama el teorema de Bolzano-Weierstrass. Establece que para cualquier conjunto infinito acotado de puntos existe un punto tal que en toda vecindad de él hay puntos del conjunto.

Cauchy había usado sin prueba (en una de sus demostraciones de la existencia de raíces de un polinomio) la existencia de un mínimo de una función continua definida sobre un intervalo cerrado. Weierstrass, en sus clases en Berlín, demostró que para cualquier función continua de una o más variables definida sobre un dominio cerrado y acotado existe un valor mínimo y un valor máximo de la función.

En su trabajo inspirado por las ideas de Georg Cantor y Weierstrass, Heine definió la continuidad uniforme para funciones de una o varias variables [12] y después demostró [13] que una función que es continua en un intervalo cerrado y acotado de los números reales es

[12] *Jour. für Math.*, 71, 1870, 353-365.
[13] *Jour. für Math.*, 74, 1872, 172-188.

uniformemente continua. El método de Heine introdujo y usó el siguiente teorema: sean dados un intervalo cerrado $[a,b]$ y un conjunto infinito numerable Δ de intervalos, todos en $[a,b]$, tales que todo punto x de $a \leqslant x \leqslant b$ sea un punto interior de al menos uno de los intervalos de Δ. (Los extremos a y b se consideran como puntos interiores cuando a es el extremo izquierdo de un intervalo y b el extremo derecho de otro intervalo.) Entonces un conjunto que consiste en un número *finito* de intervalos de Δ tiene la misma propiedad, a saber, todo punto del intervalo cerrado $[a,b]$ es un punto interior de al menos uno de los intervalos de este conjunto finito (a y b pueden ser extremos).

Emile Borel (1871-1956), uno de los grandes matemáticos franceses de este siglo, reconoció la importancia de poder seleccionar un número finito de intervalos recubridores, y lo estableció primero como un teorema independiente para el caso de un conjunto original de intervalos Δ numerable.[14] Aunque muchos matemáticos alemanes y franceses se refieren a este teorema como el teorema de Borel, ya que Heine usó en su demostración la propiedad de la continuidad uniforme, el teorema también es conocido como teorema de Heine-Borel. El mérito del teorema, como lo señaló Lebesgue, no radica en su demostración, la cual no es difícil, sino en percibir su importancia y enunciarlo como un teorema distinto. El teorema se aplica a conjuntos cerrados en cualquier número de dimensiones y es ahora básico en la teoría de los conjuntos.

La extensión del teorema de Heine-Borel para el caso en que puede seleccionarse un conjunto finito de intervalos recubridores de un conjunto infinito no numerable, se le atribuye generalmente a Lebesgue, quien reclamó haber conocido el teorema en 1898 y haberlo publicado en sus *Leçons sur l'intégration* (Lecciones sobre la integración, 1904). Sin embargo, fue publicado primero por Pierre Cousin (1867-1933) en 1895.[15]

3. La derivada

D'Alembert fue el primero en ver que Newton tenía esencialmente la noción correcta de la derivada. D'Alembert dice explícita-

---

[14] *Ann. de l'Ecole Norm. Sup.* (3), 12, 1895, 9-55.
[15] *Acta Math.*, 19, 1895, 1-61.

mente en la *Encyclopédie* que la derivada debe basarse en el límite de la razón de las diferencias de variables dependientes e independientes. Esta versión es una reformulación de las razones primera y última de Newton. D'Alembert no avanzó más porque sus pensamientos aún estaban atados a la intuición geométrica. Sus sucesores de los cincuenta años siguientes todavía fracasaron en dar una definición clara de la derivada. Hasta Poisson creía que hay números positivos que no son cero, pero que son más pequeños que cualquier número dado por muy pequeño que sea.

Bolzano fue el primero (1817) en definir la derivada de $f(x)$ como la cantidad $f'(x)$ a la cual se acerca indefinidamente la razón $[f(x + \Delta x) - f(x)]/\Delta x$ conforme Δx se acerca a 0 con valores positivos y negativos. Bolzano insistió en que $f'(x)$ no era un cociente de ceros o una razón de cantidades evanescentes sino un número al cual tendía la razón anterior.

En su *Résumé des leçons* (Resumen de Lecciones),[16] Cauchy definió la derivada de la misma manera que Bolzano. Entonces unificó esta noción con la de diferenciales de Leibniz definiendo dx como cualquier cantidad finita y dy como $f'(x)dx$.[17] En otras palabras, se introducen dos cantidades dx y dy cuya razón, por definición, es $f'(x)$. Las diferenciales tienen significado en términos de la derivada y son meramente una noción auxiliar de la cual se podría prescindir lógicamente, pero son convenientes como una manera de pensar o escribir. Cauchy también señaló lo que significaban las expresiones diferenciales utilizadas durante el siglo XVIII en términos de derivadas.

A continuación clarificó la relación entre $\Delta y/\Delta x$ y $f'(x)$ a través del teorema del valor medio, esto es, $\Delta y = f'(x + \theta\Delta x)\Delta x$, donde $0 < \theta < 1$. El teorema mismo era conocido por Lagrange (cap. 20, sec. 7). La demostración de Cauchy del teorema del valor medio usaba la continuidad de $f'(x)$ en el intervalo Δx.

Aunque Bolzano y Cauchy habían rigorizado (de alguna manera) las nociones de continuidad y de derivada, Cauchy y casi todos los matemáticos de su época creyeron, y muchos textos «demostraron» durante los siguientes cincuenta años, que una función continua debe ser diferenciable (excepto desde luego en puntos aislados tales como $x = 0$ para $y = 1/x$). Bolzano sí entendió la distinción entre conti-

[16] 1823, *Œuvres* (2), 4, 22.

[17] Lacroix ya había definido de esta manera *dy* en la primera edición de su *Traité*.

nuidad y diferenciabilidad. En su *Funktionenlehre* (Lecciones de Funciones), que escribió en 1834, pero que no terminó ni publicó,[18] dio un ejemplo de una función continua que no tiene derivada finita en ningún punto. El ejemplo de Bolzano, al igual que sus otros trabajos, pasó desapercibido.[19] Aún si hubiese sido publicado en 1834, probablemente no hubiera causado impresión, porque la curva no tenía una representación analítica, y para los matemáticos de ese período las funciones todavía eran entidades dadas por expresiones analíticas.

El ejemplo que finalmente hizo ver de manera cabal la distinción entre continuidad y diferenciabilidad fue dado por Riemann en el *Habilitationsschrif*, el artículo de 1854 que escribió para calificarse como *Privatdozent* en Göttingen, «Uber die Darstellbarkeit einer Function durch eine trigonometrische Reihe» («Sobre la representabilidad de una función mediante una serie trigonométrica»).[20] (El artículo sobre los fundamentos de la geometría (cap. 37, sec. 3) fue dado como una *conferencia* de concurso.) Riemann definió la siguiente función: denótese con (x) la diferencia entre x y el entero más cercano y sea $(x) = 0$ si está a la mitad entre dos enteros consecutivos. Entonces $-1/2 < (x) < 1/2$. Ahora $f(x)$ se define como

$$f(x) = \frac{(x)}{1} + \frac{(2x)}{4} + \frac{(3x)}{9} + \dots$$

Esta serie converge para todos los valores de x. Sin embargo, para $x = \varrho/2n$, donde ϱ es un entero primo con $2n$, $f(x)$ es discontinua y tiene un salto cuyo valor es $\pi^2/8n^2$. En todos los otros valores de x, $f(x)$ es continua. Más aún, $f(x)$ es discontinua un número infinito de veces en todo intervalo arbitrariamente pequeño. Sin embargo, $f(x)$ es integrable (sec. 4). Lo que es más, $F(x) = \int f(x)dx$ es continua para toda x, pero no tiene derivada donde $f(x)$ es discontinua. Esta función patológica no atrajo mucha atención hasta que fue publicada en 1868.

---

[18] *Schriften*, 1, Praga, 1930. Fue editado y publicado por K. Rychlik, Praga, 1930.

[19] En 1922 Rychlik demostró que la función era no diferenciable en todas partes. Véase Gerhard Kowalewski, «Uber Bolzanos nichtdifferenzierbare stetige Funktion», *Acta Math.*, 44, 1923, 315-319. Este artículo contiene una descripción de la función de Bolzano.

[20] *Abh. der Ges. der Wiss. zu Gött.*, 13, 1868, 87-132 = *Werke*, 227-264.

Una distinción aún más sorprendente entre continuidad y diferenciabilidad fue mostrada por el matemático suizo Charles Cellerier (1818-1889). En 1860 dio un ejemplo de una función que es continua pero en ningún sitio diferenciable, a saber,

$$f(x) = \sum_{n=1}^{\infty} a^{-n} \operatorname{sen} a^{n}x,$$

en la cual a es un entero positivo grande. Sin embargo, esto no se publicó hasta 1890.[21] El ejemplo que atrajo la mayor atención se debe a Weierstrass. Remontándonos hasta 1861, había afirmado en sus conferencias que cualquier intento por demostrar que la diferenciabilidad se sigue de la continuidad debe ser fallido. Dio entonces, en una conferencia a la Academia de Berlín el 18 de julio de 1872, el ejemplo clásico de una función continua en ninguna parte diferenciable.[22] Weierstrass comunicó su ejemplo en una carta de 1874 a Du Bois-Reymond y el ejemplo fue publicado primero por este último.[23] La función de Weierstrass es

$$f(x) = \sum_{n=0}^{\infty} b^{n} \cos (a^{n}\pi x)$$

donde a es un entero impar y b una constante positiva menor que 1 tal que $ab > 1 + (3\pi/2)$. La serie es uniformemente convergente y, por tanto, define una función continua. El ejemplo proporcionado por Weierstrass apresuró la creación de muchas más funciones que son continuas en un intervalo o en todas partes pero que no son diferenciables ya sea en un conjunto denso de puntos o en cualquier punto.[24]

La significación histórica del descubrimiento de que la continuidad no implica la diferenciabilidad y de que las funciones pueden tener todos los tipos de comportamiento anormal fue grandiosa. Hizo que los matemáticos fueran más aprensivos al confiar en la intuición o en el pensamiento geométrico.

---

[21] *Bull. des Sci. Math.* (2), 14, 1890, 142-160.

[22] *Werke*, 2, 71-74.

[23] *Jour. für Math.*, 79, 1875, 21-37.

[24] Otros ejemplos y referencias se pueden encontrar en E. J. Townsend, *Functions of Real Variables*, Henry Holt, 1928, y en E. W. Hobson, *The Theory of Functions of a Real Variable*, 2, cap. 6, Dover (reimp.), 1957.

4. La integral

La obra de Newton mostró cómo se podía calcular áreas invirtiendo la diferenciación. Desde luego que, aun ahora, este es el método esencial. La idea de Leibniz de área y volumen como una «suma» de elementos tales como rectángulos o cilindros [la integral definida] fue desdeñada. Cuando se empleó a fondo este último concepto en el siglo XVIII se usó de manera imprecisa.

Cauchy hizo hincapié en definir la integral como el límite de una suma, en lugar de considerarla como la inversa de la diferenciación. Hubo al menos una razón principal para el cambio. Fourier, como sabemos, trató con funciones discontinuas, y la fórmula para los coeficientes de una serie de Fourier, a saber,

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\mathrm{sen}\,nx\,dx$$

necesita de las integrales de tales funciones. Fourier consideró la integral como una suma (el punto de vista leibniziano) y así no tuvo dificultad para manejar incluso $f(x)$ discontinuas. Sin embargo, se tuvo que considerar el problema del significado analítico de la integral cuando $f(x)$ es discontinua.

El ataque más sistemático de Cauchy sobre la integral definida fue hecho en su *Résumé* (1823), en donde también señala que es necesario establecer la existencia de la integral definida e indirectamente de la antiderivada o función primitiva antes de que puedan usarse. Comienza con funciones continuas.

Da la definición precisa de la integral para $f(x)$ continuas [25] como el límite de una suma. Si el intervalo $[x,X]$ se subdivide por los valores de x: x_1, x_1, ..., x_{n-1}, con $x_n = X$, entonces la integral es

$$\lim_{n\to\infty}\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}),$$

donde ξ_i es cualquier valor de x en $[x_{i-1},x_i]$. La definición presupone que $f(x)$ es continua sobre $[x_0,X]$ y que la longitud del subintervalo más grande tiende a cero. La definición es aritmética. Cauchy de-

[25] *Résumé*, 81-84 = *Œuvres* (2), 4, 122-127.

muestra que la integral existe, no importa cómo se escojan los x_i y ξ_i. Sin embargo, su prueba no era rigurosa porque no poseía la noción de continuidad uniforme. Denota el límite usando la notación propuesta por Fourier $\int_{x_0}^{X} f(x)dx$ en lugar de $\int f(x)\,dx \begin{bmatrix} x = b \\ x = a \end{bmatrix}$ empleada frecuentemente por Euler para la antidiferenciación.

Cauchy define entonces

$$F(x) = \int_{x_0}^{x} f(x)dx$$

y muestra que $F(x)$ es continua en $[x_0, X]$. Formando

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_{x}^{x+h} f(x)\,dx$$

y usando el teorema del valor medio para integrales, Cauchy prueba que

$$F'(x) = f(x).$$

Este es el teorema fundamental del cálculo, y la presentación que hace Cauchy es la primera demostración de él. Entonces, después de demostrar que todas las primitivas de una $f(x)$ dada difieren en una constante, define la integral indefinida como

$$\int f(x)dx = \int_{a}^{x} f(x)dx + C,$$

y señala que

$$\int_{a}^{b} f'(x) = f(b) - f(a)$$

presupone que $f'(x)$ es continua. Después, Cauchy trata las integrales singulares (impropias), donde $f(x)$ se hace infinita en algún valor de x en el intervalo de integración, o donde el intervalo de integración

se extiende hasta ∞. Para el caso en que $f(x)$ tiene una discontinuidad en $x = c$ en cuyo valor $f(x)$ puede ser acotada o no, Cauchy define

$$\int_a^b f(x)dx = \lim_{\varepsilon_1 \to 0} \int_a^{c-\varepsilon_1} f(x)dx + \lim_{\varepsilon_2 \to 0} \int_{c+\varepsilon_2}^b f(x)dx$$

cuando estos límites existen. Cuando $\varepsilon_1 = \varepsilon_2$ obtenemos lo que Cauchy llamó el valor principal.

Las nociones de área acotada por una curva, longitud de una curva, volumen acotado por superficies y áreas de superficies se habían aceptado como entendidas intuitivamente, y se había considerado uno de los grandes logros del cálculo que estas cantidades pudiesen calcularse por medio de integrales. Pero Cauchy, teniendo presente su meta de aritmetizar el análisis, *definió* estas cantidades geométricas por medio de las integrales que habían sido formuladas para calcularlas. Cauchy, involuntariamente, impuso una limitación en los conceptos que definió debido a que las fórmulas del cálculo imponen restricciones sobre las cantidades involucradas. Así, la fórmula para la longitud de arco de una curva dada por $y = f(x)$ es

$$s = \int_a^b \sqrt{1 + (y')^2}\, dx$$

y esta fórmula presupone la diferenciabilidad de $f(x)$. Posteriormente iba a surgir la cuestión de cuáles son las definiciones más generales de áreas, longitudes de curvas, y volúmenes (cap. 42, sec. 5).

Cauchy había demostrado la existencia de una integral para cualquier integrando continuo. También había definido la integral cuando el integrando tiene discontinuidades de salto e infinitas. Pero con el desarrollo del análisis se hizo manifiesta la necesidad de considerar integrales de funciones de comportamiento más irregular. El tema de la integrabilidad fue retomado por Riemann en su artículo en 1854 sobre series trigonométricas. Dice que, al menos para las matemáticas, es importante, aunque no para las aplicaciones físicas, considerar las condiciones más amplias bajo las cuales se cumple la fórmula integral para los coeficientes de Fourier.

Riemann generalizó la integral para abarcar funciones $f(x)$ definidas y acotadas en un intervalo $[a,b]$. Descompone este intervalo

en subintervalos[26] $\Delta x_1, \Delta x_2, ..., \Delta x_n$ y define la oscilación de $f(x)$ en Δx cómo la diferencia entre el valor más grande y el más pequeño de $f(x)$ en Δx_i. Entonces demuestra que una condición necesaria y suficiente para que las sumas

$$S = \sum_{i=1}^{n} f(x_i)\, \Delta x_i,$$

donde x_i es cualquier valor de x en Δx_i, tiendan a un único límite (que la integral exista) cuando el máximo Δx_i tiende a cero es que la suma de los intervalos Δx_i en los que la oscilación de $f(x)$ es más grande que cualquier número dado λ debe tender a cero con el tamaño de los intervalos.

Riemann señala entonces que esta condición sobre las oscilaciones le permite reemplazar las funciones continuas por funciones con discontinuidades aisladas y también por funciones que tengan un conjunto denso en todas partes de puntos de discontinuidad. De hecho, el ejemplo que dio de una función integrable con un número infinito de discontinuidades en todo intervalo arbitrariamente pequeño (sec. 3) lo ofreció para ilustrar la generalidad de su concepto de integral. Así, Riemann prescindió de la continuidad y continuidad a trozos en la definición de la integral.

En su artículo de 1854, Riemann, sin observaciones adicionales, proporciona otra condición necesaria y suficiente para que una función acotada $f(x)$ sea integrable en $[a,b]$. Requiere que primero se establezca lo que ahora se llaman las sumas superior e inferior

$$S = M_1 \Delta x_1 + ... + M_n \Delta x_n$$
$$s = m_1 \Delta x_1 + ... + m_n \Delta x_n$$

donde m_i y M_i son los valores mínimo y máximo de $f(x)$ en Δx_i. Después, haciendo $D_i = M_i - m_i$, Riemann afirma que la integral de $f(x)$ en $[a,b]$ existe si y sólo si

$$\lim_{\max \Delta x \to 0} \{D_1\, \Delta x_1 + D_2\, \Delta x_2 + ... + D_n\, \Delta x_n\} = 0$$

para todas las elecciones de Δx_i que llenen el intervalo $[a,b]$. Darboux completó esta formulación y demostró que la condición es

[26] Por brevedad se usa Δx_i para los subintervalos y sus longitudes.

necesaria y suficiente.[27] Existen muchos valores de S correspondiendo cada uno de ellos a una partición de $[a,b]$ en Δx_i. De manera análoga, existen muchos valores de s. Cada S se llama una suma superior y cada s una suma inferior. Sea J la máxima cota inferior de las S e I la mínima cota superior de las s. Se sigue que $I \leqslant J$. El teorema de Darboux establece entonces que las sumas S y s tienden respectivamente a J e I cuando se incrementa indefinidamente el número de Δx_i de tal manera que el máximo subintervalo tiende a cero. Se dice que una función acotada es integrable en $[a,b]$ si $J = I$.

Darboux demuestra a continuación que una función acotada será integrable sobre $[a,b]$ si y sólo si las discontinuidades de $f(x)$ constituyen un conjunto de medida cero. Por esto último quería dar a entender que los puntos de discontinuidad pueden encerrarse en un conjunto finito de intervalos cuya longitud total es arbitrariamente pequeña. Esta misma formulación de la condición de integrabilidad fue establecida por otros autores en el mismo año (1875). Los términos integral superior y la notación $\overline{\int}_a^b f(x)\, dx$ para la máxima cota inferior J de las S, e integral inferior y la notación $\underline{\int}_a^b f(x)\, dx$ para la mínima cota superior I de las s fueron introducidos por Volterra.[28]

Darboux también demostró en el artículo de 1875 que el teorema fundamental del cálculo se cumple para funciones integrables en el sentido ampliado. Bonnet había dado una demostración del teorema del valor medio del cálculo diferencial que no utilizaba la continuidad de $f'(x)$.[29] Darboux, usando esta demostración, la cual es ahora habitual, demostró que

$$\int_a^b f'(x)dx = f(b) - f(a)$$

cuando f' es integrable solamente en el sentido Riemann-Darboux. El argumento de Darboux fue que

$$f(b) - f(a) = \sum_{i=1}^{n} f(x_i) - f(x_{i-1}),$$

[27] *Ann. de l'Ecole Norm. Sup.* (2), 4, 1875, 57-112.
[28] *Gior. di Mat.*, 19, 1881, 333-372.
[29] Publicado en el *Cours de calcul différentiel et intégral* de Serret, 1, 1868, 17-19.

donde $a = x_0 < x_1 < x_2 < ... < x_n = b$. Por el teorema del valor medio

$$\sum f(x_i) - f(x_{i-1}) = \sum f'(t_i)(x_i - x_{i-1}),$$

donde t_i es algún valor en (x_{i-1}, x_i). Ahora, si el máximo Δx_i, o $x_i - x_{i-1}$, tiende a cero, entonces el miembro de la derecha de esta última ecuación tiende a $\int_a^b f'(x)dx$ y el miembro izquierdo es $f(b) - f(a)$.

Una de las actividades favoritas de los setenta y ochenta fue construir funciones con diversos conjuntos infinitos de discontinuidades que aún así fueran integrables en el sentido de Riemann. En este sentido, H. J. S. Smith [30] dio el primer ejemplo de una función no integrable en el sentido de Riemann, pero para la cual los puntos de discontinuidad eran «raros». La función de Dirichlet (sec. 2) también es no integrable en este sentido, pero es discontinua en todas partes.

La noción de integración se extendió entonces a funciones no acotadas y a varias integrales impropias. La extensión más significativa fue llevada a cabo en el siguiente siglo por Lebesgue (cap. 44). Sin embargo, para 1875, en lo que concernía al cálculo elemental, la noción de integral era suficientemente amplia y estaba rigurosamente fundamentada.

También se abordó la teoría de las integrales dobles. Los casos más sencillos habían sido tratados en el siglo XVIII (cap. 19, sec. 6). En su artículo de 1814 (cap. 27, sec. 4), Cauchy demostró que el orden de integración en el que se calcula una integral doble $\iint f(x,y)\, dx\, dy$ sí es importante si el integrando es discontinuo en el dominio de integración. Cauchy señaló específicamente [31] que las integrales repetidas

$$\int_0^1 dy\left(\int_0^1 f(x, y)\, dx\right), \int_0^1 dx\left(\int_0^1 f(x, y)\, dy\right)$$

no necesitan ser iguales cuando f es no acotada.

Karl J. Thomae (1840-1921) extendió la teoría de Riemann sobre

---

[30] *Proc. Lon. Math. Soc.*, 6, 1875, 140-153 = *Coll. Papers*, 2, 86-100.

[31] *Mémoire* de 1814; véase, en particular, p. 394 de *Œuvres* (1), 1.

integración a funciones de dos variables.[32] Entonces Thomae, en 1878,[33] proporcionó un ejemplo sencillo de una función acotada para la cual la segunda integral repetida existe pero la primera carece de significado.

En los ejemplos de Cauchy y Thomae la integral doble no existe. Pero en 1883,[34] Du Bois-Reymond demostró que, aun cuando la integral doble exista, las dos integrales repetidas no necesariamente existen. En el caso de las integrales dobles también la generalización más significativa fue realizada por Lebesgue.

5. La teoría de series infinitas

Los matemáticos del siglo XVIII usaron las series indiscriminadamente. A finales del siglo, algunos resultados dudosos o llanamente absurdos del trabajo con series infinitas estimularon las investigaciones sobre la validez de las operaciones con ellas. Alrededor de 1810, Fourier, Gauss y Bolzano iniciaron el manejo exacto de las series infinitas. Bolzano enfatizó que se debe considerar la convergencia y, en particular, criticó la demostración imprecisa del teorema del binomio. Abel fue el crítico más franco de los usos antiguos de las series.

En su artículo de 1811, y su *Teoría analítica del calor,* Fourier dio una definición satisfactoria de la convergencia de una serie infinita, aunque en general trabajó libremente con series divergentes. En el libro (pág. 196 de la edición inglesa) describe la convergencia dando a entender que conforme n se incrementa, la suma de n términos tiende a un valor fijo de manera cada vez más próxima y debiera diferir de él solamente por una cantidad que se haga menor que cualquier magnitud dada. Aún más, reconoció que la convergencia de una serie de funciones puede obtenerse solamente en un intervalo de valores x. También enfatizó que una condición necesaria para la convergencia es que los términos tiendan a cero. Sin embargo, la serie $1 - 1 + 1 + \ldots$ aún lo confundía; tomó como su suma $1/2$.

La primera investigación importante y estrictamente rigurosa de la convergencia fue hecha por Gauss en su artículo de 1812 «Dis-

---

[32] *Zeit. für Math. und Phys.*, 21, 1876, 224-227.
[33] *Zeit. für Math. und Phys.*, 23, 1878, 67-68.
[34] *Jour. für Math.*, 94, 1883, 273-290.

quisitions Generales Circa Seriem Infinitam» (Investigaciones generales de las series infinitas) [35] en donde estudió la serie hipergeométrica $F(\alpha,\beta,\gamma,\chi)$. En la mayor parte de su obra llamó convergente a una serie si los términos a partir de un α determinado decrecen hacia cero. Pero en su artículo de 1812 hizo notar que este no es el concepto correcto. Debido a que la serie hipergeométrica puede representar a muchas funciones para diferentes elecciones de α, β y γ, le parecía deseable desarrollar un criterio exacto para la convergencia de esta serie. Se llega a ese criterio de manera laboriosa, pero clarifica la cuestión de la convergencia para los casos que fue destinada a cubrir. Demostró que la serie hipergeométrica converge para x real y compleja si $|x| < 1$ y diverge si $|x| > 1$. Para $x = 1$, la serie converge si y solo si $\alpha + \beta < \gamma$ y para $x = -1$ la serie converge si y sólo si $\alpha + \beta < \gamma + 1$. El rigor inusual desanimó el interés de los matemáticos de la época por el artículo. Lo que es más, Gauss tenía interés en las series particulares y no consideró los principios generales de la convergencia de las series.

Aunque Gauss es mencionado frecuentemente como uno de los primeros en reconocer la necesidad de restringir el uso de las series a sus dominios de convergencia, evitó cualquier posición definitiva. Estaba tan interesado en resolver problemas concretos mediante cálculos numéricos que hizo uso del desarrollo de Stirling de la función gamma. Cuando investigó la convergencia de la serie hipergeométrica en 1812 indicó,[36] que lo hizo así, no para dar su propia postura sobre el tema, sino para complacer a aquellos que favorecían el rigor de los geómetras antiguos. En la presentación de su artículo [37] usó el desarrollo de $\log(2 - 2\cos x)$ en cosenos de múltiplos de x a pesar de que no había demostración de la convergencia de esta serie y no podía haber demostración con las técnicas disponibles en la época. En su obra astronómica y geodésica, Gauss, al igual que los matemáticos del siglo XVIII, siguió la práctica de usar un número finito de términos de una serie infinita y desdeñar los demás. Dejaba de incluir términos cuando veía que los términos siguientes eran numéricamente pequeños y, por supuesto, no estimaba el error.

Poisson tomó una posición particular. Rechazó las series diver-

---

[35] *Comm. Soc. Gött.*, 2, 1813 = *Werke*, 3, 125-162 y 207-229.
[36] *Werke*, 3, 129.
[37] *Werke*, 3, 156.

gentes[38] y hasta dio ejemplos de cómo los cálculos con series divergentes pueden conducir a resultados falsos. Pero de todas maneras hizo uso extensivo de las series divergentes en su representación de funciones arbitrarias por medio de series de funciones trigonométricas y esféricas.

Bolzano, en su publicación de 1817, tenía la noción correcta de la condición para la convergencia de una sucesión, la condición ahora atribuida a Cauchy. Bolzano también tenía nociones claras y correctas acerca de la convergencia de las series. Pero, como ya hemos notado, su obra no llegó a ser ampliamente conocida.

El trabajo de Cauchy sobre la convergencia de series es el primer tratamiento extensivo significativo del tema. En su *Cours d'analyse* Cauchy dice: «Sea

$$s_n = u_0 + u_1 + u^2 + \dots + u_{n-1}$$

la suma de los n primeros términos [de la serie infinita que se considera], designando n un número natural. Si, para valores constantemente crecientes de n, la suma s_n se acerca indefinidamente a cierto límite s, la serie es llamada *convergente*, y el límite en cuestión será llamado la *suma* de la serie.[39] Por el contrario, si mientras n se incrementa indefinidamente, la suma s_n no se acerca a un límite fijo, la serie será llamada *divergente* y no tendrá suma.

Después de definir la convergencia y la divergencia, Cauchy establece (*Cours*, p. 125) su criterio para la convergencia, a saber, una sucesión (S_n) converge a un límite S si y sólo si $S_{n+r} - S_n$ puede hacerse en valor absoluto menor que cualquier cantidad asignable para todo r y n suficientemente grandes. Cauchy demuestra que esta condición es necesaria pero únicamente dice que si la condición se satisface, se asegura la convergencia de la sucesión. Carecía del conocimiento de las propiedades de los números reales para hacer la demostración.

Cauchy establece enseguida y demuestra casos específicos para la convergencia de series con términos positivos. Señala que u_n debe tender a cero. Otra prueba (*Cours*, 132-135) requiere que se encuentre el límite o límites hacia los cuales la expresión $(u_n)^{1/n}$ tiende con-

---

[38] *Jour. de l'Ecole Poly.*, 19, 1823, 404-509.

[39] La noción correcta del límite de una sucesión fue dada por Wallis en 1695 (*Opera*, 1695, 1, 382), pero no fue adoptada.

forme n se hace infinito, y designa el mayor de estos límites por k. Entonces la serie será convergente si $k < 1$ y divergente si $k > 1$. También proporciona la prueba del cociente que emplea $\lim_n u_{n+1}/u_n$. Si este límite es menor que 1, la serie converge y si es mayor que 1, la serie diverge. Se proporcionan casos especiales si la razón es 1. Siguen criterios de comparación y un criterio logarítmico. Demuestra que la suma $u_n + v_n$ de dos series convergentes converge a la suma de las sumas separadas y el resultado análogo para el producto. Las series con algunos términos negativos, demuestra Cauchy, convergen cuando la serie de los valores absolutos de los términos converge, y después deduce la prueba de Leibniz para series alternadas.

Cauchy también considera la suma de una serie

$$\sum u_n(x) = u_1(x) + u_2(x) + u_3(x) + \dots$$

en la que todos los términos son funciones continuas reales unívocas. Los teoremas sobre la convergencia de las series de términos constantes se aplican aquí para determinar un intervalo de convergencia. También considera series cuyos términos sean funciones complejas.

Lagrange fue el primero en enunciar el teorema de Taylor con resto, pero Cauchy, en sus textos de 1823 y 1829 hizo la observación importante de que la serie de Taylor infinita converge a la función de la cual se obtiene si el residuo tiende a cero. Proporciona el ejemplo $e^{-x^2} + e^{-1/x^2}$ de una función cuya serie de Taylor no converge a la función. En su texto de 1823, da el ejemplo e^{-1/x^2} de una función que tiene todas las derivadas en $x = 0$ pero no tiene expansión de Taylor alrededor de $x = 0$. Aquí, por medio de un ejemplo, contradice la afirmación de Lagrange en su *Théorie des Fonctions* (cap. V, art. 30) de que si $f(x)$ tiene todas las derivadas en x_0, entonces puede expresarse como una serie de Taylor que converge a $f(x)$ para x cercana a x_0. Cauchy también dio [40] una forma alternativa para el residuo en la fórmula de Taylor.

Aquí cometió Cauchy algunos errores adicionales con respecto al rigor. En su *Cours d'analyse* (pp. 131-132) establece que $F(x)$ es continua si cuando $F(x) = \Sigma_1^\infty u_n(x)$ la serie es convergente y las

[40] *Exercices de mathématiques*, 1, 1826. 5 = *Œuvres*, (2), 6, 38-42.

$u_n(x)$ son continuas. En su *Résumé des leçons*,[41] dice que si las $u_n(x)$ son continuas y la serie converge, entonces la serie se puede integrar término a término; esto es,

$$\int_a^b F\,dx = \sum_1^\infty \int_a^b u_n dx.$$

Pasó por alto la necesidad de la convergencia uniforme. También afirma que para funciones continuas [42]

$$\frac{\partial}{\partial u}\int_a^b f(x, u)\,dx = \int_a^b \frac{\partial f}{\partial u}\,dx.$$

La obra de Cauchy inspiró a Abel. Escribiendo de París a su primer maestro Holmboë en 1826, Abel dijo [43] que Cauchy «es en el presente quien conoce cómo deben ser tratadas las matemáticas». En ese año,[44] Abel investigó el dominio de convergencia de la serie binómica

$$1 + mx + \frac{m(m-1)}{2}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \ldots$$

con m y x complejos, y expresó su admiración de que nadie hubiese investigado previamente la convergencia de esta importantísima serie. Primero demuestra que si la serie

$$f(\alpha) = v_0 + v_1\alpha + V_2\alpha^2 + \ldots,$$

en donde las v_i son constantes y α es real, converge para un valor δ de α entonces convergerá para todo valor de α más pequeño, y $f(\alpha - \beta)$ para β acercándose a 0 tenderá a $f(\alpha)$ cuando α sea igual o menor que δ. La última parte dice que una serie de *potencias* convergente es una función continua de su argumento hasta δ inclusive, pues α puede ser δ.

En este mismo artículo de 1826,[45] Abel corrigió el error de

---

[41] 1823, *Œuvres* (2), 4, p. 237.
[42] *Exercices de mathématiques*, 2, 1827 = *Œuvres* (2), 7, 160.
[43] *Œuvres*, 2, 259.
[44] *Jour. für Math.*, 1, 1826, 311-339 = *Œuvres*, 1, 219-250.
[45] *Œuvres*, 1, 224.

Cauchy sobre la continuidad de la suma de una serie convergente de funciones continuas. Dio el ejemplo de

$$\operatorname{sen} x - \frac{\operatorname{sen} 2x}{2} + \frac{\operatorname{sen} 3x}{3} \dots \qquad (2)$$

que es discontinua cuando $x = (2n + 1)\,\pi$ y n es entero, aunque los términos individuales son continuos.[46] Entonces, usando la *idea* de convergencia uniforme dio una demostración correcta de que la suma de una serie uniformemente convergente de funciones continuas es continua en el interior del intervalo de convergencia. Abel no aisló la propiedad de la convergencia uniforme de una serie.

La noción de convergencia uniforme de una serie $\Sigma_1^\infty\, u_n(x)$ requiere que dado cualquier ε, exista un N tal que para todo $n > N$, $|S(x) - \Sigma_1^\infty\, u_n(x)| < \varepsilon$ para todo x en algún intervalo. $S(x)$ es por supuesto la suma de la serie. Esta noción fue reconocida en y por sí misma por Stokes, un físico matemático sobresaliente,[47] e independientemente por Philipp L. Seidel (1821-1896).[48] Ninguno de los dos proporcionó la formulación precisa. Más bien ambos demostraron que si la suma de una serie de funciones continuas es discontinua en x_0 entonces existen valores de x cercanos a x_0 para los cuales la serie converge de manera arbitrariamente lenta. Tampoco relacionaron la necesidad de la convergencia uniforme para la justificación de integrar una serie término a término. De hecho, Stokes aceptó [49] el uso de Cauchy de la integración término a término. Finalmente, Cauchy reconoció la necesidad de la convergencia uniforme [50] para asegurar la continuidad de la suma de una serie de funciones continuas pero, incluso él, en esa época, no vio el error en el uso de la integración de una serie término a término.

De hecho, Weierstrass [51] poseía la noción de convergencia uniforme tan temprano como en 1842. En un teorema que, sin saberlo,

---

[46] La serie (2) es la expansión de Fourier de $x/2$ en el intervalo $-\pi < x < \pi$. Por tanto, la serie representa la función periódica que vale $x/2$ en cada intervalo de longitud 2π. Entonces la serie converge a $\pi/2$ cuando x tiende a $(2n + 1)\,\pi$ por la izquierda y la serie converge a $-\pi/2$ cuando x tiende a $(2n + 1)$ por la derecha.

[47] *Trans. Camb. Phil. Soc.*, 8, 1848, 533-583 = *Math. and Phys. Papers*, 1, 236-313.

[48] *Abh. der Bayer. Akad. der Wiss.*, 1847/1849, 379-394.

[49] *Papers*, 1, 242, 255, 268 y 283.

[50] *Comp. Rend.*, 36, 1853, 454-459 = *Œuvres* (1), 12, 30-36.

[51] *Werke*, 1, 67-85.

duplica el teorema de Cauchy sobre la existencia de las soluciones en series de potencias de un sistema de ecuaciones diferenciales ordinarias de primer orden, afirma que las series convergen uniformemente y así constituyen funciones analíticas de la variable compleja. Aproximadamente por la misma época, Weierstrass usó la noción de convergencia uniforme para proporcionar condiciones para la integración de una serie término a término y condiciones para la diferenciación bajo el signo de integral.

A través del círculo de estudiantes de Weierstrass se dio a conocer la importancia de la convergencia uniforme. Heine enfatizó la noción en un artículo sobre series trigonométricas.[52] Es posible que Heine haya aprendido la idea por medio de Georg Cantor, quien había estudiado en Berlín y después fue en 1867 a Halle, donde Heine era profesor de matemáticas.

Durante sus años como maestro de escuela de educación media, Weierstrass también descubrió que cualquier función continua sobre un intervalo cerrado del eje real se puede expresar en ese intervalo como una serie de polinomios absoluta y uniformemente convergente. Weierstrass incluyó también funciones de varias variables. Este resultado [53] despertó considerable interés y se establecieron muchas extensiones de él para la representación de funciones complejas por medio de una serie de polinomios o una serie de funciones racionales en el último cuarto del siglo XIX.

Se había supuesto que los términos de una serie se pueden ordenar como se desee. En un artículo de 1837,[54] Dirichlet demostró que en una serie absolutamente convergente es posible agrupar o reacomodar los términos y no cambiar la suma. También proporcionó ejemplos para mostrar que los términos de cualquier serie condicionalmente convergente se pueden ordenar de modo que la suma se altere. Riemann, en un artículo escrito en 1854 (véase más adelante), demostró que por medio de reordenaciones adecuadas de los términos, la suma podía ser igual a cualquier número dado. Los grandes matemáticos desarrollaron muchos más criterios para la convergencia de series infinitas durante los años treinta y todo el resto del siglo.

---

[52] *Jour. für Math.*, 71, 1870, 353-365.

[53] *Sitzungsbber. Akad. Wiss. zu Berlin*, 1855, 633-639, 789-905 = *Werke*, 3, 1-37.

[54] *Abh. König. Akad. der Wiss., Berlin*, 1837, 45-81 = *Werke*, 1, 313-342 = *Jour. de Math.*, 4, 1839, 393-422.

6. Las series de Fourier

Como sabemos, el trabajo de Fourier demostró que una gran variedad de funciones puede ser representada por medio de series trigonométricas. Quedó abierto el problema de encontrar condiciones precisas para las funciones que poseen una serie de Fourier convergente. Los esfuerzos de Cauchy y Poisson fueron infructuosos.

Dirichlet se interesó por las series de Fourier después de conocer a Fourier en París durante los años 1822-1825. En un artículo básico, «Sur la convergence des séries trigonométriques» (Sobre la convergencia de las series trigonométricas),[55] Dirichlet dio el primer conjunto de condiciones *suficientes* para que la serie de Fourier representando una $f(x)$ dada converja y lo haga a $f(x)$. La demostración dada por Dirichlet es un refinamiento de la que bosquejó Fourier en las secciones finales de su *Teoría analítica del calor.* Considérese $f(x)$ dada, ya sea periódica con período 2π o dada en el intervalo $[-\pi,\pi]$, y definida periódica en cada intervalo de longitud 2π a la izquierda y a la derecha de $[-\pi,\pi]$. Las condiciones de Dirichlet son:

a) $f(x)$ es unívoca y acotada.

b) $f(x)$ es continua a trozos; esto es, sólo tiene un número finito de discontinuidades en el período (cerrado).

c) $f(x)$ es monótona a trozos; esto es, sólo tiene un número finito de máximos y mínimos en un período.

La $f(x)$ puede tener diferentes representaciones analíticas en partes diferentes del período fundamental.

El método de demostración de Dirichlet consistía en hacer una sumatoria directa de n términos e investigar lo que sucede conforme n tiende a infinito. Demostró que para cualquier valor dado de x la suma de la serie es $f(x)$ siempre que $f(x)$ sea continua en ese valor de x y es $(1/2)[f(x-0)+f(x+0)]$ si $f(x)$ es discontinua en ese valor de x.

En su demostración, Dirichlet se vio forzado a discutir cuidadosamente los valores límites de las integrales

$$\int_0^a f(x)\,\frac{\operatorname{sen}\mu x}{\operatorname{sen}x}\,dx, \qquad a>0$$

[55] *Jour. für Math.*, 4, 1829, 157-169 = *Werke*, 1, 117-132.

$$\int_a^b f(x)\,\frac{\operatorname{sen}\mu x}{\operatorname{sen} x}\,dx, \qquad b > a > 0$$

conforme μ crece indefinidamente. Aún se les llama a éstas integrales de Dirichlet.

Fue en relación con este trabajo cuando proporcionó la función que vale c para valores racionales de x y d para valores irracionales de x (sec. 2). Había esperado generalizar la noción de integral de modo que aún fuera posible representar una clase más amplia de funciones por medio de una serie de Fourier convergente a estas funciones, pero la función particular dada se pensó como un ejemplo de una función que no pudiera ser incluida en una noción más amplia de integral.

Riemann estudió durante un tiempo bajo la dirección de Dirichlet en Berlín y adquirió interés por las series de Fourier. En 1854 emprendió la investigación del tema en su *Habilitationsschrift* en Göttingen,[56] «Uber die Darstellbarkeit einer Funnction durch eine trigonometrische Reihe», que tenía como finalidad encontrar condiciones necesarias y suficientes que una función debe satisfacer para que la serie de Fourier para $f(x)$ en un punto x del intervalo $[-\pi,\pi]$ convergiera a $f(x)$.

Riemann sí demostró el teorema fundamental de que si $f(x)$ es acotada e integrable en $[-\pi,\pi]$ entonces la coeficientes de Fourier

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx \qquad (3)$$

se acercan a cero cuando n tiende al infinito. El teorema también demostraba que para $f(x)$ acotada e integrable, la convergencia de su serie de Fourier en un punto en $[-\pi,\pi]$ depende solamente del comportamiento de $f(x)$ en la vecindad de ese punto. Sin embargo, el problema de encontrar condiciones necesarias y suficientes para $f(x)$ de modo que su serie de Fourier converja a $f(x)$ no se había y no ha sido resuelto.

Riemann abrió otra línea de investigación. Consideró las series *trigonométricas* pero no exigió que los coeficientes se determinaran

[56] *Abh. der Ges. der Wiss. zu Gött.*, 13, 1868, 87-132 = *Werke*, 227-264.

por la fórmula (3) para los coeficientes de Fourier. Comienza con la serie

$$\sum_{1}^{\infty} a_n \operatorname{sen} nx + \frac{b_0}{2} + \sum_{1}^{\infty} b_n \cos nx \qquad (4)$$

y define

$$A_0 = (1/2)b_0, \qquad A_n(x) = a_n sen\ nx + b_n cos\ nx.$$

Entonces la serie (4) es igual a

$$f(x) = \sum_{n=0}^{\infty} A_n(x).$$

Desde luego que $f(x)$ tiene un valor solamente para aquellos valores de x para los cuales la serie converge. Refirámonos a la serie misma por Ω. Ahora los términos de Ω pueden aproximarse a cero para toda x o para alguna x. Estos dos casos los trata Riemann por separado.

Si a_n y b_n tienden a cero, los términos de Ω tienden a cero para toda x. Sea $F(x)$ la función

$$F(x) = C + C'x + A_0 \frac{x^2}{2} - A_1 - \frac{A_2}{4} - \dots - \frac{A_n}{n^2} \dots$$

que se obtiene por medio de dos integraciones sucesivas de Ω. Riemann demuestra que $F(x)$ converge para toda x y es continua en x. Entonces la propia $f(x)$ puede integrarse. Riemann demuestra luego una serie de teoremas acerca de $F(x)$, que como consecuencia conducen a condiciones necesarias y suficientes para que una serie de la forma (4) converja a una función dada $f(x)$ de período 2π. Proporciona, entonces, una condición necesaria y suficiente para que la serie trigonométrica (4) converja en un valor particular de x, con a_n y b_n aproximándose todavía a cero conforme n tiende a ∞.

Después considera el caso alterno donde $\lim_{n\to\infty} A_n$ depende del valor de x y da condiciones que se cumplen cuando la serie Ω es convergente para valores particulares de x y un criterio para la convergencia en valores particulares de x.

También demuestra que una $f(x)$ dada puede ser integrable y aún así no tener una representación en serie de Fourier. Lo que es más,

existen funciones no integrables a las cuales converge la serie Ω para un número infinito de valores de x tomados entre límites arbitrariamente cercanos. Finalmente, una serie trigonométrica puede converger para un número infinito de valores de x en un intervalo arbitrariamente pequeño a pesar de que a_n y b_n no tiendan a cero para todas las x.

La naturaleza de la convergencia de las series de Fourier recibió mayor atención después de la introducción del concepto de convergencia uniforme por Stokes y Seidel. Desde Dirichlet se sabía que las series, en general, eran sólo condicionalmente convergentes, si es que lo eran, y que su convergencia dependía de la presencia de términos positivos y negativos. Heine hizo notar, en un artículo de 1870,[57] que la demostración usual de que una $f(x)$ acotada está representada unívocamente entre $-\pi$ y π por una serie de Fourier es incompleta porque la serie puede no ser uniformemente convergente y así no puede integrarse término a término. Esto sugirió que, sin embargo, pueden existir series trigonométricas no uniformemente convergentes que sí representen una función. Además, una función continua podría representarse por una serie de Fourier y aún así la serie podría no ser uniformemente convergente. Estos problemas dieron origen a la aparición de una nueva serie de investigaciones que buscaban establecer la unicidad de la representación de una función mediante una serie trigonométrica y si los coeficientes son necesariamente los coeficientes de Fourier. Heine demostró, en el artículo que se mencionó antes, que una serie de Fourier que representa una función acotada que satisface las condiciones de Dirichlet es uniformemente convergente en las partes del intervalo $[-\pi,\pi]$ que quedan cuando se eliminan del intervalo entornos arbitrariamente pequeños de los puntos de discontinuidad de la función. En estos entornos la convergencia es necesariamente no uniforme. Heine demostró después que si la convergencia uniforme que se acaba de especificar se cumple para una serie trigonométrica que representa a una función, entonces la serie es única.

El segundo resultado, sobre unicidad, es equivalente a la afirmación de que si una serie trigonométrica de la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx) \qquad (5)$$

[57] *Jour. für Math.*, 71, 1870, 353-365.

es uniformemente convergente y representa cero donde converge, esto es, excepto en un conjunto finito P de puntos, entonces los coeficientes son todos cero y por supuesto que entonces la serie representa a cero en todo el intervalo $[-\pi,\pi]$.

Los problemas asociados con la unicidad de las series trigonométricas y las de Fourier atrajeron a Georg Cantor, quien estudió el trabajo de Heine. Cantor empezó sus investigaciones buscando criterios de unicidad para las representaciones de las funciones en series trigonométricas. Demostró [58] que cuando $f(x)$ se representa por una serie trigonométrica convergente para toda x, no existe otra serie trigonométrica de la misma forma que converja análogamente para todo x y represente la misma función $f(x)$. Otro artículo [59] proporcionó una demostración mejor de este último resultado.

El teorema de unicidad que demostró puede enunciarse de nuevo así: si, para toda x, existe una representación convergente de cero por una serie trigonométrica, entonces los coeficientes a_n y b_n son cero. Entonces Cantor demuestra, en el artículo de 1871, que la conclusión es válida aun si se prescinde de la convergencia para un número finito de valores de x. Este artículo fue el primero de una serie de ellos en los que Cantor trata los conjuntos de valores excepcionales de x. Extendió [60] el resultado de unicidad al caso donde se permite un conjunto infinito de valores excepcionales. Para describir este conjunto definió primero que un punto p es punto límite de un conjunto de puntos S si todo intervalo que contenga a p contiene infinitos puntos de S. Después introdujo la noción de conjunto derivado de un conjunto de puntos. Este conjunto derivado consiste de los puntos límite del conjunto original. Existe entonces un segundo conjunto derivado, esto es, el conjunto derivado del conjunto derivado, y así sucesivamente. Si el n-ésimo conjunto derivado de un conjunto dado es un conjunto finito de puntos entonces se dice que el conjunto dado es de clase n-ésima o n-ésimo orden (o de primera especie). La respuesta final de Cantor a la cuestión de si una función puede tener dos representaciones diferentes en series trigonométricas en el intervalo $[-\pi,\pi]$ o si cero puede tener una representación de Fourier que no sea cero, es que si en el intervalo una serie trigonométrica suma cero para toda x excepto las de un

---

[58] *Jour. für Math.*, 72, 1870, 139-142 = *Ges. Abh.*, 80-83.
[59] *Jour. für Math.*, 73, 1871, 294-296 = *Ges. Abh.*, 84-86.
[60] *Math. Ann.*, 5, 1872, 123-132 = *Ges. Abh.*, 92-102.

conjunto de puntos de clase n-ésima (en el cual no se sabe algo más acerca de la serie) entonces todos los coeficientes de la serie deben ser cero. En este artículo de 1872, Cantor sentó las bases de la teoría de los conjuntos de puntos que consideraremos en un capítulo posterior. Muchos otros investigaron el problema de la unicidad en la última parte del siglo XIX y la primera del XX.[61]

Durante aproximadamente cincuenta años después del trabajo de Dirichlet, se creyó que la serie de Fourier de cualquier función continua en $[-\pi,\pi]$ converge a la función. Pero Du Bois-Reymond[62] dio un ejemplo de una función continua en $(-\pi,\pi)$ cuya serie de Fourier no converge en un punto particular. También construyó otra función continua cuya serie de Fourier no converge en los puntos de un conjunto denso en todas partes. Después, en 1875,[63] demostró que si una serie trigonométrica de la forma

$$a_0 + \sum_1^\infty (a_n \cos nx + b_n \operatorname{sen} nx)$$

converge a $f(x)$ en $[-\pi,\pi]$ y si $f(x)$ es integrable (en un sentido aún más general que el de Riemann, en el que $f(x)$ puede ser no acotada en un conjunto de primera especie) entonces la serie debe ser la serie de Fourier para $f(x)$. También demostró[64] que cualquier serie de Fourier de una función que es integrable en el sentido de Riemann se puede integrar término a término a pesar de que la serie no sea uniformemente convergente.

Muchos autores emprendieron entonces la investigación del problema ya resuelto de una manera por Dirichlet, a saber, proporcionar condiciones suficientes para que una función $f(x)$ tenga una serie de Fourier convergente a $f(x)$. Varios resultados son clásicos. Jordan proporcionó una condición suficiente en términos del concepto de función de variación acotada, que él introdujo.[65] Sea $f(x)$ acotada en $[a,b]$ y sea $a = x_0, x_1, \ldots, x_{n-1}, x_n = b$ un modo de división (parti-

---

[61] Se puede encontrar detalles en E. W. Hobson, *The Theory of Functions of a Real Variable*, vol. 2, 656-698.

[62] *Nachrichten König. Ges. der Wiss. zu Gött.*, 1873, 571-582.

[63] *Abh. der Bayer. Akad. der Wiss.*, 12, 1876, 117-166.

[64] *Math. Ann.*, 22, 1883, 260-268.

[65] *Comp. Rend.*, 92, 1881, 228-230 = *Œuvres*, 4, 393-395 y *Cours d'analyse*, 2, 1.ª ed., 1882, cap. V.

ción) de este intervalo. Sean $y_0, y_1, ..., y_{n-1}, y_n$ los valores de $f(x)$ en estos puntos. Entonces, para cada partición

$$\sum_{0}^{n-1} (y_{r+1} - y_r) = f(b) - f(a)$$

denótese con t a

$$\sum_{0}^{n-1} |y_{r+1} - y_r|.$$

Para cada modo de subdividir el intervalo $[a,b]$ existe una t. Cuando, correspondiendo a todos los modos posibles de división de $[a,b]$, las sumas t tienen una mínima cota superior, entonces f se define como de variación acotada en $[a,b]$.

La condición suficiente de Jordan establece que la serie de Fourier para la función integrable $f(x)$ converge a

$$(1/2)\ [f(x + 0) + f(x - 0)]$$

en cada punto para el cual exista un entorno en el que $f(x)$ sea de variación acotada.[66]

Durante los sesenta y setenta se examinaron también las propiedades de los coeficientes de Fourier y entre los resultados importantes obtenidos estaba el que es llamado teorema de Parseval (quien lo enunció bajo condiciones más restringidas, cap. 29, sec. 3), según el cual si $f(x)$ y $[f(x)]^2$ son integrables en el sentido de Riemann en $[-\pi,\pi]$ entonces

$$\frac{1}{\pi}\int_{-\pi}^{\pi} [f(x)]^2 dx = 2a_0^2 + \sum_{1}^{\infty} (a_n^2 + b_n^2),$$

y si $f(x)$ y $g(x)$ y sus cuadrados son integrables en el sentido de Riemann entonces

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\, dx = 2a_0\alpha_0 + \sum_{1}^{\infty} (a_n\alpha_n + b_n\beta_n),$$

[66] *Cours d'analyse*, 2.ª ed., 1893, 1, 67-72.

donde a_n, b_n, α_n y β_n son los coeficientes de Fourier de $f(x)$ y $g(x)$ respectivamente *.

7. La situación del análisis

La obra de Bolzano, Cauchy, Weierstrass y otros proporcionó rigor al análisis. Esta obra liberó al cálculo y sus extensiones de toda dependencia de nociones geométricas, del movimiento y de comprensiones intuitivas. Desde el principio, estas investigaciones causaron considerable conmoción. Después de una reunión científica en la que Cauchy presentó la teoría sobre la convergencia de las series, Laplace se marchó precipitadamente a casa y permaneció allí retirado hasta que hubo examinado las series en su *Mécanique celeste*. Afortunadamente, encontró que todas eran convergentes. Cuando se llegó a conocer la obra de Weierstrass a través de sus conferencias, el efecto fue aún más notable. Las mejoras en el rigor pueden verse si se compara la primera edición del *Cours d'analyse* de Jordan (1882-1887) con la segunda (1893-1896) y la tercera (3 vols., 1909-1915). Muchos otros tratados incorporaron el nuevo rigor.

La rigorización del análisis no resultó ser el fin de la investigación en los fundamentos. Esto fue así porque prácticamente todo el trabajo presuponía el sistema de los números reales, pero este tema permanecía desorganizado. Excepto para Weierstrass, quien, como veremos, consideró el problema del número irracional durante los años cuarenta, todos los demás no creían necesario investigar los fundamentos lógicos del sistema de los números. Parecería que aún los más grandes matemáticos deben desarrollar sus capacidades para apreciar la necesidad del rigor en etapas. El trabajo sobre los fundamentos lógicos del sistema de los números reales iba a seguir en breve (cap. 41).

El descubrimiento de que las funciones continuas no tienen ne-

---

* En 1966 L. Carleson demostró que la serie de Fourier de una función de cuadrado integrable en el sentido de Lebesgue converge en casi todo punto a la función, vid. L. Carleson, «On convergence and growth of partial sums of Fourier series», *Acta Mathematica*, 116 (1966), 135-157. Un poco después, Hunt extendió el resultado usando los métodos de Carleson, a L^p con $p \neq 2$, vid. R. A. Hunt, «On the convergence of Fourier series», en *Orthogonal expansions and their continuous analogues*, Proc. Conference Southern Illinois University, 1967; Southern Illinois University Press, 1968, pp. 235-255. (Nota del traductor.)

cesariamente derivadas, que las funciones discontinuas pueden integrarse, la nueva luz arrojada sobre las funciones discontinuas por el trabajo de Dirichlet y Riemann acerca de las series de Fourier y el estudio de la variedad y extensión de las discontinuidades de las funciones hicieron que los matemáticos se dieran cuenta de que el estudio riguroso de las funciones se extiende más allá de las que se usan en el cálculo y las ramas usuales del análisis, donde el requerimiento de la diferenciabilidad restringe generalmente la clase de funciones. El estudio de las funciones continuó en el siglo XX y llevó al desarrollo de una nueva rama de las matemáticas, conocida como la teoría de las funciones de una variable real (cap. 44).

Igual que todos los nuevos movimientos en las matemáticas, la rigorización del análisis no avanzó sin oposición. Hubo mucha controversia en cuanto a si los refinamientos en el análisis debían investigarse. Las funciones peculiares que se introdujeron fueron atacadas como curiosidades, funciones disparatadas, funciones divertidas y como juguetes matemáticos, tal vez más intrincados pero no de mayor consecuencia que los cuadrados mágicos. También se vieron como enfermedades o parte de la patología mórbida de las funciones y sin tener relación alguna con los problemas importantes de las matemáticas puras y aplicadas. Estas nuevas funciones, violando leyes que se creía perfectas, se consideraron como signos de anarquía y caos que se mofaban del orden y armonía que las generaciones previas habían buscado. Las muchas hipótesis que ahora se tenían que hacer para enunciar un teorema preciso fueron vistas como pedantes y destructoras de la elegancia del análisis clásico del siglo XVIII, «como era en el paraíso», por parafrasear a Du Bois-Reymond. Los nuevos detalles se resintieron como oscurecedores de las ideas principales.

Poincaré, en particular, desconfió de esta nueva investigación. Dijo que:[67]

> La lógica a veces crea monstruos. Durante medio siglo, hemos visto una masa de funciones extrañas que parecen forzadas a parecerse tan poco como sea posible a funciones honestas que sirvan a algún propósito. Más continuidad, o menos continuidad, más derivadas, etc. Ciertamente, desde el punto de vista de la lógica, estas funciones extrañas son las más generales; por otro lado, aquellas que uno se encuentra sin buscarlas, y que siguen

[67] *L'Enseignement mathématique*, 11, 1899, 157-162 = *Œuvres*, 2, 129-134.

leyes sencillas se presentan como un caso particular que no representa más que una esquina pequeña.

En épocas pasadas, cuando uno inventaba una nueva función era con un propósito práctico; hoy, uno las inventa con el propósito de mostrar defectos en los razonamientos de nuestros padres y sólo eso se deducirá de ellas.

Charles Hermite dijo en una carta a Stieltjes: «Me aparto con miedo y horror de esta lamentable plaga de funciones que no tienen derivadas.»

Du Bois-Reymond [68] dio a conocer otro tipo de objeción. Su interés radicaba en que la aritmetización del análisis separaba al análisis de la geometría, y consecuentemente de la intuición y el pensamiento físico. Reducía el análisis «a un simple juego de símbolos donde los signos escritos toman la significación arbitraria de las piezas en el ajedrez o un juego de cartas».

El asunto que provocó la mayor controversia fue la prohibición de las series divergentes, principalmente por Abel y Cauchy. En una carta a Holmboë, escrita en 1826, Abel dice: [69]

Las series divergentes son la invención del demonio, y es una vergüenza basar en ellas cualquier demostración, sea la que fuere. Usándolas, se puede obtener cualquier conclusión que le plazca a uno y por eso han producido tantas falacias y tantas paradojas... He llegado a estar prodigiosamente atento a todo esto, pues con la excepción de las series geométricas, no existe en todas las matemáticas una sola serie infinita cuya suma se haya determinado rigurosamente. En otras palabras, las cosas más importantes en matemáticas son aquellas que tienen menor fundamentación.

Sin embargo, Abel mostró cierto interés acerca de si se había pasado por alto alguna idea buena, porque continúa en su carta del siguiente modo: «Que la mayor parte de estas cosas son correctas a pesar de eso es extraordinariamente sorprendente. Estoy tratando de encontrar una razón de ello; es una cuestión extraordinariamente interesante.» Abel murió joven y, por tanto, nunca investigó el asunto.

Cauchy, también, tuvo algunos escrúpulos en condenar a las series divergentes al ostracismo. Dice en la introducción de su *Cours* (1821), «He sido forzado a admitir diversas proposiciones que pa-

---

[68] *Théorie générale des fonctions*, 1887, 61.
[69] *Œuvres*, 2, 256.

recen algo deplorables, por ejemplo, que una serie divergente no puede sumarse.» A pesar de esta conclusión, Cauchy continuó usando series divergentes, según aparece en notas añadidas a la publicación en 1827 [70] de un ensayo para concursar, escrito en 1815, sobre ondas de agua. Decidió investigar la cuestión de por qué las series divergentes resultaron tan útiles, y de hecho finalmente llegó a estar cerca de reconocer la razón (cap. 47).

Los matemáticos franceses aceptaron la prohibición de Cauchy de las series divergentes, pero no los ingleses y alemanes. En Inglaterra, la escuela de Cambridge defendió el uso de las series divergentes apelando al principio de la permanencia de la forma (cap. 32, sec. 1). En relación con las series divergentes, el principio fue usado primero por Robert Woodhouse (1773-1827). En *The Principles of Analytic Calculation (Los principios del Cálculo Analítico,* 1803, p. 3) señala que en la ecuación

$$\frac{1}{1-r} = 1 + r + r^2 + \ldots \tag{6}$$

el signo de igualdad tiene «un significado más amplio» que sólo el de la igualdad numérica. Por tanto la ecuación se cumple tanto si la serie es divergente como si no.

Peacock también aplicó el principio de permanencia de la forma a las operaciones con series divergentes.[71] En la página 267 dice: «Así, como para $r < 1$, la igualdad (6) anterior se cumple, entonces para $r = 1$ obtenemos $\infty = 1 + 1 + 1 + \ldots$ Para $r > 1$ obtenemos un número negativo a la izquierda y, como los términos a la derecha se incrementan continuamente, una cantidad mayor que infinito en la derecha.» Peacock acepta esto. El punto que trata de aclarar es que la serie pueda representar $1/(1 - r)$ para toda r. Dice,

Si las operaciones del álgebra se consideran como generales, y los símbolos que se sujetan a ellas ilimitados en valor, será imposible evitar la formación de series divergentes al igual que convergentes; y si tales series se consideran como los resultados de operaciones que son definibles, además de las series mismas, entonces no será muy importante entrar en tal examen

[70] *Mém. des sav. étrangers,* 1, 1827, 3-312; véase *Œuvres* (1), 1, 238, 277, 286.
[71] *Report on the Recent Progress and Present State of Certain Branches of Analysis,* Brit. Assn. for Adv. of Science, 3, 1833, 185-352.

de la relación de los valores aritméticos de los términos sucesivos como puede ser necesario para asegurar su convergencia o divergencia; pues bajo tales circunstancias, se deben considerar como formas equivalentes que representan a su función generadora, y como que para los propósitos de tales operaciones poseen propiedades equivalentes... El intento por excluir el uso de las series divergentes en las operaciones simbólicas necesariamente impondría un límite sobre la universalidad de las fórmulas y operaciones algebraicas, lo cual, conjuntamente, es contrario al espíritu de la ciencia... Necesariamente conduciría a una mayor y embarazosa multiplicación de casos: privaría a casi todas las operaciones algebraicas de mucha de su certeza y simplicidad.

Augustus de Morgan, aunque mucho más agudo y más consciente que Peacock de las dificultades con las series divergentes, se encontraba, sin embargo, bajo la influencia de la escuela inglesa y también se impresionó por los resultados obtenidos por el uso de las series divergentes a pesar de las dificultades en ellas. En 1844 empezó un artículo, incisivo pero aun así confuso, sobre «Series Divergentes»,[72] con estas palabras, «Creo que será generalmente admitido que el encabezamiento de este artículo describe el único asunto que aún queda, de carácter elemental, sobre el cual existe un serio cisma entre los matemáticos en cuanto a la exactitud o inexactitud absolutas de los resultados.» La posición que De Morgan tomó ya la había declarado en su *Differential and Integral Calculus (Cálculo Diferencial e Integral)*,[73] «La historia del álgebra nos muestra que nada es más erróneo que el rechazo de cualquier método que surja naturalmente, a causa de uno o más casos aparentemente válidos en los cuales tal método conduzca a resultados equivocados. Tales casos ciertamente debieran enseñar cautela, pero no rechazo; si se hubiera preferido lo último a lo primero, las cantidades negativas, y aún más sus raíces cuadradas, habrían sido un obstáculo eficaz para el progreso del álgebra... y esos campos inmensos del análisis en los cuales incluso los que rechazaban las series divergentes ahora se incluyen sin temor, no habrían sido descubiertos, ni mucho menos cultivados y establecidos... La consigna que yo adoptaría contra un desarrollo que me parece calculado para detener el progreso del descubrimiento estaría contenido en una palabra y un símbolo —recuérdese $\sqrt{-1}$.» Distingue entre el significado aritmético y algebraico de una serie.

---

[72] *Trans. Camb. Philo. Soc.*, parte II, 1844, 182-203, pub. 1849.
[73] Londres, 1842, p. 566.

El significado algebraico se cumple en todos los casos. Para dar cuenta de algunas de las conclusiones falsas obtenidas con las series divergentes dice en el artículo de 1844 (p. 187) que la integración es una operación aritmética y no algebraica y por tanto no se podría aplicar sin razonamiento adicional con las series divergentes. Pero la obtención de

$$\frac{1}{1-r} = 1 + r + r^2 + \ldots$$

comenzando con $y = 1 + ry$, reemplazando y en la derecha por $1 + ry$, y continuando así, lo acepta porque es algebraico. Análogamente, de $z = 1 + 2z$ se obtiene $z = 1 + 2 + 4 + \ldots$ Por tanto $-1 = 1 + 2 + 4 + \ldots$ y esto es correcto. Acepta la teoría completa (como estaba en esa época) de las series trigonométricas pero estaría dispuesto a rechazarla si se pudiera dar un ejemplo donde $1 - 1 + 1 - 1 + 1 \ldots$ no sea igual a 1/2 (véase cap. 20).

Muchos otros matemáticos ingleses prominentes proporcionaron otros tipos de justificación para la aceptación de las series divergentes, algunos retornando a un argumento de Nicholas Bernoulli (cap. 20, sec. 7), de que la serie (6) contiene un resto r^∞ o $r^\infty/(1 - r^\infty)$. Esto debe tomarse en cuenta (aunque no indicaron cómo). Otros dijeron que la suma de una serie divergente es algebraicamente verdadera pero aritméticamente falsa.

Algunos matemáticos alemanes usaron los mismos argumentos que Peacock aunque utilizaron palabras distintas, tales como operaciones sintácticas opuestas a operaciones ariméticas o literal opuesto a numérico. Martin Ohm [74] dijo, «Una serie infinita (soslayando cualquier cuestión sobre la convergencia o divergencia) está completamente adaptada para representar una expresión dada si uno puede estar seguro de tener la ley correcta para desarrollar la serie. Del *valor* de una serie infinita se puede hablar sólo si converge.» Los argumentos en Alemania a favor de la legitimidad de las series divergentes fueron sostenidos durante varias décadas más.

La defensa del uso de las series divergentes no estuvo tan cercana del desastre como podría parecer, aunque muchos de los argumentos dados a favor de las series eran, tal vez, inverosímiles. Por un lado,

---

[74] *Aufsätze aus dem Gebiet der höheren Mathematik (Ensayos en el Dominio de las Matemáticas Avanzadas,* 1823).

en la totalidad del análisis del siglo XVIII, la atención al rigor o a la demostración fue mínima, y esto era aceptable porque los resultados obtenidos casi siempre eran correctos. Así, los matemáticos llegaron a acostumbrarse a los procedimientos y argumentos imprecisos. Pero, aún más interesante para la cuestión, muchos de los conceptos y operaciones que habían causado perplejidades, tales como los números complejos, se demostró que eran correctos después de que fueran totalmente entendidos. Por eso los matemáticos pensaron que las dificultades con la series divergentes también serían aclaradas cuando se tuviera una mejor comprensión, y que las series divergentes resultarían ser legítimas. Lo que es más, las operaciones con las series divergentes estaban ligadas con frecuencia con otras operaciones poco entendidas del análisis, tales como el intercambio del orden de los límites, la integración sobre discontinuidades de un integrando y la integración sobre un intervalo infinito, de modo que los defensores de las series divergentes podían sostener que las conclusiones falsas atribuidas al uso de las series divergentes surgieron de otras fuentes de problemas.

Un argumento que pudo haberse esgrimido es que cuando una función analítica se expresa en algún dominio por una serie de potencias, lo que Weierstrass llamó un elemento, esta serie ciertamente sí lleva consigo las propiedades «algebraicas» o «sintácticas» de la función y estas propiedades son llevadas más allá del dominio de convergencia del elemento. El proceso de la continuación analítica usa este hecho. Realmente había material matemático sólido en el concepto de las series divergentes que justificaba su utilidad. Pero el reconocimiento de este material y la aceptación final de las series divergentes tuvo que esperar una nueva teoría de las series infinitas (cap. 47).

Bibliografía

Abel, N. H.: *Œuvres complètes*, 2 vols., 1881, Johnson Reprint Corp., 1964.

— *Mémorial publié a l'occasion du centenaire de sa naissance*, Jacob Dybwad, 1902. Cartas escritas por Abel y dirigidas a él.

Bolzano, B.: *Rein analytischer Beweis des Lehrsatzes, das zwischen je zweit Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reele Wurzel der Gleichung liege*, Gottlieb Hass, Prague, 1817 = *Abh. Königl. Böhm. Ges. der Wiss.* (3), 5, 1814-1817, pub. 1818 = *Ostwald's*

Klassiker der exakten Wissenschaften, 153, 1905, 3-45. No contenida en los *Schriften* de Bolzano.
— *Paradoxes of the Infinite,* Routledge and Kegan Paul, 1950. Contiene una investigación histórica del trabajo de Bolzano.
— *Schriften*, 5 vols., Königlichen Böhmischen Gesellschaft der Wissenschaften, 1930-1948.
Boyer, Carl B.: *The Concepts of the Calculus,* Dover (reimp.), 1949, cap. 7.
Burkhardt, H.: «Uber den Gebrauch divergenter Reihen in der Zeit von 1750-1860.» *Math. Ann.,* 70, 1911, 169-206.
— «Trigonometrische Reihe und Integrale». *Encyk. der Math. Wiss.,* II A12, 819-1354, B. G. Teubner, 1904-1916.
Cantor, Georg: *Gesammelte Abhandlungen* (1932), Georg Olms (reimp.), 1962.
Cauchy, A. L.: *Œuvres* (2), Gauthier-Villars, 1897-1899, vols. 3 y 4.
Dauben, J. W.: «The Trigonometric Background to Georg Cantor's Theory of Sets». *Archive for History of Exact Sciences,* 7, 1971, 181-216.
Dirichlet, P. G. L.: *Werke,* 2 vols. Georg Reimer, 1889-1897, Chelsea (reimp.), 1969.
Du Bois-Reymond, Paul: *Zwei Abhandlungen über unendliche und trigonometrische Reihen* (1871 y 1874), Ostwald's Klassiker, 185; Wilhelm Engelmann, 1913.
Freudenthal, H.: «Did Cauchy Plagiarize Bolzano?». *Archive for History of Exact Sciences,* 7, 1971, 375-392.
Gibson, G. A.: «On the History of Fourier Series». *Proc. Edindburgh Math. Soc.,* 11, 1892-1893, 137-166.
Grattan-Guinness, I.: «Bolzano, Cauchy and the "New Analysis" of the Nineteenth Century». *Archive for History of Exact Sciences,* 6, 1970, 372-400.
— *The Development of the Foundations of Mathematical Analysis from Euler to Riemann,* Massachusetts Institute of Technology Press, 1970.
Hawkins, Thomas W., Jr.: *Lebesgue's Theory of Integration: Its Origins and Development.* University of Wisconsin Press, 1970, caps. 1-3.
Manheim, Jerome H.: *The Genesis of Point Set Topology.* Macmillan, 1964, caps. 1-4.
Pesin, Ivan H.: *Classical and Modern Integration Theories.* Academic Press, 1970, cap. 1.
Pringsheim, A.: «Irrationalzahlen und Konvergenz unendlichen Prozesse». *Encyk. der Math. Wiss.,* IA3, 47-147, B. G. Teubner, 1898-1904.
Reiff, R.: *Geschichte der unendlichen Reihen.* H. Lauppsche Bbuchhandlung, 1889; Martin Sänding (reimp.), 1969.
Riemann, Bernhard: *Gesammelte mathematische Werke,* 2.ª ed. (1902), Dover (reimp.), 1953.
Schleinger, L.: «Uber Gauss Arbeiten zur Funktionenlehre». *Nachrichten*

König. Ges. der Wiss. zu Gött., 1912, Beiheft, 1-43. También en el *Werke* de Gauss, 10, 77 y sigs.

Schoenflies, Arthur M.: «Die Entwicklung der Lehre von den Punktmannigfaltigkeiten». *Jahres. der Deut. Math.-Verein.*, 8, 1899, 1-250.

Singh, A. N.: «The Theory and Construction of Non-Differentiable Functions», en E. W. Hobson: *Squaring the Circle and Other Monographs*, Chelsea (reimp.), 1953.

Smith, David E.: *A Source Book in Mathematics*, Dover (reimp.), 1959, vol. 1, 286-291, vol. 2, 635-637.

Stolz, O.: «B. Bolzanos Bedeutung in der Geschichte der Infinitesimalrechnung». *Math. Ann.*, 18, 1881, 255-279.

Weierstrass, Karl: *Mathematische Werke*, 7 vols., Mayer und Muller, 1894-1927.

Young, Grace C.: «On Infinite Derivatives». *Quart. Jour. of Math.*, 47, 1916, 127-175.

Capítulo 41

LA FUNDAMENTACION DE LOS NUMEROS REALES Y TRANSFINITOS

Dios hizo los enteros; el resto es obra del hombre.

LEOPOLD KRONECKER

1. Introducción

Uno de los hechos más sorprendentes en la historia de la matemática es que no se acometiera la fundamentación lógica del sistema de números reales hasta finales del siglo XIX. Hasta ese momento no quedaron lógicamente establecidas ni siquiera las propiedades más simples de los números racionales positivos y negativos y de los números irracionales, ni habían sido definidos esos números. La misma fundamentación lógica de los números complejos, cuya existencia no databa de mucho antes (vid. cap. 32, sec. 1), presuponía la del sistema de números reales. Considerando el extenso desarrollo del álgebra y el análisis, y cómo en ellos se utilizan los números reales, la falta de una estructuración precisa de éstos y de sus propiedades muestra cuán ilógicamente progresa la matemática. La comprensión intuitiva de esos números parecía suficiente, y los matemáticos se contentaban con operar sobre esa base.

La rigorización del análisis impelía a remediar la falta de claridad en el sistema numérico mismo. Por ejemplo, la demostración de Bolzano (cap. 40, sec. 2) de que una función continua que es negativa para $x = a$ y positiva para $x = b$ se anula para algún valor de x

entre *a* y *b* patinaba en un punto crítico porque faltaba una adecuada comprensión de la estructura del sistema de números reales. El estudio detallado de los límites también mostraba la necesidad de comprender los números reales, ya que números racionales pueden tener un límite irracional y recíprocamente. La incapacidad de Cauchy para probar la suficiencia de su criterio para la convergencia de una sucesión se derivaba igualmente de su falta de comprensión de la estructura del sistema numérico. El estudio de las discontinuidades de funciones representables mediante series de Fourier revelaba la misma deficiencia. Fue Weierstrass el primero que señaló que para establecer con precisión las propiedades de las funciones continuas necesitaba la teoría del continuo aritmético.

Otra motivación para la fundamentación del sistema numérico fue el deseo de asegurar la verdad de la matemática. Como consecuencia de la creación de las geometrías no euclídeas, la geometría había perdido su *status* de verdad (vid. cap. 36, sec. 8), pero parecía todavía que la matemática construida sobre la aritmética ordinaria debía ser una realidad incuestionable en cierto sentido filosófico. Ya en 1817, en su carta a Olbers, Gauss [1] había distinguido la aritmética de la geometría en que sólo la primera era puramente a priori. En su carta a Bessel del 9 de abril de 1830,[2] repetía la afirmación de que sólo las leyes de la aritmética son necesarias y verdaderas. Sin embargo, faltaba una fundamentación del sistema numérico que despejara cualquier duda sobre la verdad de la aritmética, y del álgebra y el análisis construidos sobre esa base.

Merece la pena señalar que antes de que los matemáticos apreciaran la necesidad de analizar el sistema numérico mismo, el problema que les había parecido más pertinente era el de la fundamentación del *álgebra,* y en particular una explicación del hecho de que uno pueda usar letras para representar números reales y complejos, y operar con las letras por medio de las propiedades aceptadas como verdaderas para los enteros positivos. Para Peacock, de Morgan y Duncan Gregory, el álgebra de comienzos del siglo XIX era un ingenioso, pero también ingenuo, complejo de esquemas manipulatorios con algo de sentido pero muy poca sustancia; les parecía que el núcleo de la habitual confusión residía en una fundamentación inadecuada del álgebra. Ya hemos visto cómo resolvieron el problema

[1] *Werke,* 8, 177.
[2] *Werke,* 8, 201.

(cap. 32, sec. 1). A fines de siglo, sin embargo, quedó claro que había que profundizar por el lado del análisis y clarificar la estructura de todo el sistema numérico real. De pasada se aseguraría así también la estructura lógica del álgebra, ya que resultaba intuitivamente claro que los diferentes tipos de números poseían las mismas propiedades formales. De aquí que si se podían establecer esas propiedades sobre una base firme, podrían aplicarse a las letras que representaban números cualesquiera.

2. Números algebraicos y trascendentes

Una etapa importante hacia la mejor comprensión de los números irracionales fue el trabajo de mediados del siglo XIX sobre los irracionales algebraicos y trascendentes. La distinción entre unos y otros había quedado establecida en el siglo XVIII (cap. 25, sec. 1). El interés de tal distinción aumentó con los trabajos del XIX sobre resolución de ecuaciones, que mostraron que no todos los irracionales algebraicos pueden obtenerse mediante operaciones algebraicas sobre números racionales. Además, el problema de determinar si e y π eran algebraicos o trascendentes seguía atrayendo a los matemáticos.

Hasta 1844 siguió abierta la cuestión sobre si había o no irracionales trascendentes. Ese año, Liouville[3] mostró que cualquier número de la forma

$$\frac{a_1}{10} + \frac{a_2}{10^{2!}} + \frac{a_3}{10^{3!}} + \ldots,$$

donde los a_i son enteros arbitrarios de 0 a 9, es trascendente.

Para probarlo, Liouville demostró primero algunos teoremas sobre la aproximación de irracionales algebraicos mediante números racionales. Por definición (cap. 25, sec. 1), un número algebraico es cualquier número, real o complejo, que satisface una ecuación algebraica

$$a_0x^n + a_1x^{n-1} + \ldots + a_n = 0$$

donde los a_i son enteros. Una raíz es un número algebraico de grado

[3] *Comp. Rend.*, 18, 1844, 910-911, y *Jour. de Math.* (1), 16, 1851, 133-142.

n si satisface una ecuación de grado n y no satisface ninguna ecuación de menor grado. Algunos números algebraicos son racionales; esos son de grado uno. Liouville probó que si p/q es cualquier aproximación a un número algebraico irracional x de grado n, con p y q enteros, entonces existe un número positivo M tal que

$$\left| x - \frac{p}{q} \right| > \frac{M}{q^n}.$$

Esto significa que cualquier aproximación racional a un irracional algebraico de grado n por cualquier p/q debe ser menos precisa que M/q^n. Con otras palabras, podemos decir que si x es un irracional algebraico de grado n, existe un número positivo M tal que la desigualdad

$$\left| x - \frac{p}{q} \right| < \frac{M}{q^\mu}$$

no tiene soluciones enteras p y q para $\mu = n$ y tampoco por tanto para $\mu \leq n$. Así pues, x es trascendente si para un M fijo y para cada entero positivo μ la desigualdad tiene alguna solución p/q. Mostrando que sus irracionales satisfacen este último criterio, Liouville demostró que son trascendentes.

La siguiente gran etapa en el reconocimiento de números trascendentes específicos fue la demostración de Hermite en 1873 [4] de que e es trascendente. Después de obtener este resultado, Hermite escribió a Carl Wilhelm Borchardt (1817-1880): «No me atrevo a intentar probar la trascendencia de π. Si otros lo logran nadie estará más feliz que yo con su éxito, pero créame, mi querido amigo, que no dejará de costarles un cierto esfuerzo.»

Que π es trascendente había sido ya sospechado por Legendre (cap. 25, sec. 1). Ferdinand Lindemann (1852-1939) lo demostró en 1882 [5] con un método que no difiere esencialmente del empleado por Hermite: Lindemann probó que si $x_1, x_2, \ldots, x_n$ son números algebraicos distintos, reales o complejos, y $p_1, p_2, \ldots, p_n$ son números algebraicos no todos nulos, la suma

$$p_1 e^{x_1} + p_2 e^{x_2} + \ldots + p_n e^{x_n}$$

[4] *Comp. Rend.*, 77, 1873, 18-24, 74-79, 226-233, 285-293, y *Œuvres*, 2, 150-181.

[5] *Math. Ann.*, 20, 1882, 213-225.

no puede ser 0. Si tomamos $n = 2$, $p_1 = 1$, y $x_2 = 0$, vemos que e^{x_1} no puede ser algebraico para un x_1 que sea algebraico y no nulo. Como x_1 puede ser 1, e es trascendente. Ahora bien, dado que $e^{i\pi} + 1 = 0$, el número $i\pi$ no puede ser algebraico, y tampoco π, puesto que i lo es, y el producto de dos números algebraicos es algebraico. La demostración de que π es trascendente puso punto final a los famosos problemas de construcción de la geometría, ya que todos los números constructibles son algebraicos.

Permanece todavía el misterio en torno a una constante fundamental, la constante de Euler (cap. 20, sec. 4)

$$C = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \log n\right),$$

que es aproximadamente 0,577216 y que tiene un importante papel en análisis, en particular en el estudio de las funciones gamma y zeta, y de la que no se sabe si es racional o irracional.

3. La teoría de los números irracionales

A finales del siglo XIX se afrontó directamente la cuestión de la estructura lógica del sistema de los números reales. Los números irracionales suponían la principal dificultad. Ahora bien, el desarrollo del significado y propiedades de los números irracionales presupone la construcción del sistema de los números racionales. Los distintos autores que contribuyeron a la teoría de los números irracionales, o bien supusieron que los números racionales se conocían con tanta seguridad que no precisaban una fundamentación, o bien esbozaron precipitadamente algún esquema improvisado.

Curiosamente, la construcción de una teoría de los irracionales no requería mucho más que un cambio de punto de vista. Euclides, en el Libro V de los *Elementos,* había estudiado las razones entre magnitudes inconmensurables, y había definido la igualdad y la desigualdad entra tales razones. Su definición de igualdad (cap. 4, sec. 5) equivalía a dividir los números racionales m/n en dos clases, la de aquellos para los que m/n es menor que la razón a/b entre las magnitudes inconmensurables a y b, y la de aquellos para los que m/n es más grande. Cierto es que la lógica de Euclides era deficiente, ya que nunca definió la razón entre dos magnitudes inconmensurables.

Además, el desarrollo de Euclides de la teoría de las proporciones, o igualdades entre razones, sólo se podía aplicar a la geometría. Aun así, tuvo la idea esencial que podía haberse utilizado antes para definir los números irracionales. De hecho, Dedekind hizo uso del trabajo de Euclides y reconoció [6] su deuda hacia éste; también Weierstrass pudo guiarse por la teoría de Euclides. Sin embargo, es más fácil contemplar el pasado que el futuro. El largo retraso en sacar partido de alguna reformulación de las ideas de Euclides puede explicarse ahora fácilmente: los números negativos tenían que ser completamente aceptados antes de poder disponer del sistema completo de números racionales; además, tenía que experimentarse la necesidad de una teoría de los irracionales, y esto ocurrió solamente una vez emprendida la aritmetización del análisis.

William R. Hamilton ofreció el primer tratamiento de los números irracionales en dos artículos, leídos ante la Royal Irish Academy en 1833 y 1835,[7] y publicados más tarde como «Algebra as the Science of Pure Time». Basaba, pues, en el tiempo su noción de todos los números, racionales e irracionales, una base poco satisfactoria para las matemáticas (aunque muchos, siguiendo a Kant, la hayan considerado como una intuición básica). Tras presentar una teoría de los números racionales, introdujo la idea de separar los racionales en dos clases (describiremos más detalladamente la idea en conexión con la obra de Dedekind) y definió un número irracional como una de esas particiones, aunque no completó el trabajo.

Aparte de esta obra inacabada, todas las presentaciones de los irracionales anteriores a Weierstrass utilizaron la noción de irracional como límite de una sucesión infinita de racionales. Pero el límite, si es irracional, no tiene existencia lógica hasta que se hayan definido los irracionales. Cantor [8] señaló que este error lógico no fue advertido durante algún tiempo porque no conducía a dificultades posteriores. Weierstrass, en sus clases en Berlín a partir de 1859, ofreció una teoría de los números irracionales, señalando su necesidad. Una publicación de H. Kossak, *Die Elemente der Arithmetik* (1872), trató de presentar esta teoría, pero Weierstrass la desaprobó.

En 1869 Charles Méray (1835-1911), uno de los apóstoles de la aritmetización de las matemáticas y equivalente francés de Weiers-

---

[6] *Essays*, 40.

[7] *Trans. Royal Irish Academy*, 17, 1837, 293-422, y *Math. Papers*, 3, 3-96.

[8] *Math. Ann.*, 21, 1883, 566.

trass, dio una definición de los irracionales basada en los racionales [9]. Georg Cantor también ofreció una teoría, que necesitaba para clarificar las ideas sobre conjuntos que había utilizado en su trabajo de 1871 sobre series de Fourier. Fue seguida un año más tarde por la teoría de Heinrich Heine, que apareció en el *Journal für Mathematik*,[10] y por la de Dedekind, publicada en *Stetigkeit und irrationale Zahlen*.[11]

Todas estas teorías sobre los números irracionales son esencialmente parecidas; nos limitaremos a dar algunas indicaciones sobre las de Cantor y Dedekind. Cantor [12] comienza con los números racionales. En su artículo de 1883,[13] donde da más detalles sobre su teoría de los números irracionales, dice (p. 565) que no es necesario extenderse sobre los racionales, puesto que ya lo habían hecho Hermann Grassmann en su *Lehrbuch der Arithmetik* (1861) y J. H. T. Muller (1797-1862) en su *Lehrbuch der allgemeinen Arithmetik* (1855). De hecho, tales presentaciones no resultaron definitivas. Cantor introdujo una nueva clase de números, los números reales, que incluyen racionales e irracionales. Construyó los números reales a partir de los racionales, definiendo primeramente lo que llamaba sucesión fundamental: cualquier sucesión de racionales que cumple la condición de que, para cualquier ε prefijado, todos los términos de la sucesión, salvo a lo más un número finito, difieren entre sí en menos de ε, o bien que

$$\lim_{n\to\infty} (a_{n+m} - a_n) = 0$$

para m arbitrario. Cada sucesión fundamental es, por definición, un número real al que podemos denotar por b. Dos sucesiones fundamentales (a_n) y (b_n) son el mismo número real si y sólo si $|a_n - b_n|$ se aproxima a cero cuando n se hace infinito.

Para tales sucesiones se presentan tres posibilidades: dado un número racional arbitrario, los términos a_n de la sucesión para n suficientemente grande son todos en valor absoluto menores que el número dado; o todos los términos a partir de un cierto n son

---

[9] *Revue des Societès Savants*, 4, 1869, 280-289.
[10] *Jour. für Math.*, 74, 1872, 172-188.
[11] Continuidad y números irracionales, 1872, en *Werke*, 3, 314-334.
[12] *Math. Ann.*, 5, 1872, 123-132 = *Ges. Abh.*, 92-102.
[13] *Math. Ann.*, 21, 1883, 545-591 = *Ges. Abh.*, 165-204.

mayores que un cierto número racional positivo ϱ; o bien todos los términos de la sucesión a partir de un cierto n son menores que un cierto número racional negativo $-\varrho$. En el primer caso, $b = 0$; en el segundo $b > 0$; en el tercero $b < 0$.

Si (a_n) y (a'_n) son dos sucesiones fundamentales, denotadas por b y b', se puede probar que $(a_n \pm a_n)$ y $(a_n \cdot a_n)$ son también sucesiones fundamentales, y definen $b + b'$ y $b \cdot b'$. Además, si $b \neq 0$, entonces (a'_n/a_n) es también una sucesión fundamental, que define b'/b.

Los números reales racionales quedan incluidos en la definición de más arriba, ya que una sucesión (a_n) con todos sus términos a_n iguales al mismo número racional a define el número real racional a.

A continuación se puede definir la igualdad y desigualdad entre dos números reales cualesquiera: así, $b = b'$, $b > b'$ o $b < b'$, según que $b - b'$ sea igual, mayor o menor que 0.

El siguiente teorema es crucial. Cantor demuestra que si (b_n) es cualquier sucesión de números reales (racionales o irracionales), y si $\lim_{n\to\infty} (b_{n+m} - b_n) = 0$ para m arbitrario, entonces existe un único número real b, determinado por una sucesión fundamental (a_n) de números racionales a_n tal que

$$\lim_{n\to\infty} b_n = b.$$

Es decir, que la formación de sucesiones fundamentales de números reales no crea la necesidad de otros nuevos tipos de números que puedan servir como límites para esas sucesiones fundamentales, puesto que los números reales ya existentes bastan para proporcionar tales límites. En otras palabras, desde el punto de vista de la convergencia de sucesiones fundamentales (o lo que es lo mismo, sucesiones que satisfacen el criterio de convergencia de Cauchy), los números reales constituyen un sistema completo.

La teoría de Dedekind de los números irracionales, presentada en su libro de 1872 mencionado más arriba, parte de las ideas con que ya contaba en 1858. En esa época tuvo que dar clases de cálculo y constató que el sistema de números reales carecía de fundamentación lógica. Para probar que una cantidad monótonamente creciente y acotada se aproxima a un límite tenía que hacer uso, como otros autores, de la evidencia geométrica (decía que esa sigue siendo la forma en que hay que hacerlo en una primera presentación del cálculo, particularmente si uno no desea perder mucho tiempo). Ade-

más, muchos teoremas aritméticos básicos quedaban sin demostración; mostraba como ejemplo el hecho de que no se hubiera todavía probado rigurosamente que $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$.

En su exposición afirma que presupone el desarrollo de los números racionales, que comenta brevemente. Para introducir los irracionales se pregunta primero qué significa la continuidad geométrica. Los pensadores contemporáneos y anteriores —por ejemplo Bolzano— creían que continuidad significa la existencia de al menos otro número entre dos cualesquiera, propiedad que ahora se conoce como densidad. Pero los números racionales ellos mismos también forman un conjunto denso, luego densidad no es continuidad.

Lo que le sugirió a Dedekind la definición de número irracional fue la observación de que en cada partición de la recta en dos clases de puntos tales que cada punto de la primera está a la izquierda de cada punto de la segunda, hay *uno y sólo un punto* que produce la partición. Ese hecho hace a la recta continua. Para la recta se trata de un axioma; trasladó entonces esa idea al sistema numérico. Consideremos, dice Dedekind, cualquier partición de los números racionales en dos clases tales que cualquier número de la primera es menor que cualquier número de la segunda. Tal partición de los números racionales es lo que llama una cortadura. Si las clases son denotadas por A_1 y A_2, la cortadura se denota por (A_1,A_2). Para algunas cortaduras, específicamente las determinadas por un número racional, hay un número máximo en A_1 o un número mínimo en A_2. Recíprocamente, cada cortadura en los racionales en la que hay un máximo en la primera clase o un mínimo en la segunda está determinada por un número racional.

Pero hay cortaduras que no están determinadas por números racionales. Si ponemos en la primera clase todos los números racionales positivos y negativos cuyo cuadrado es menor que 2, y en la segunda clase todos los demás racionales, esa cortadura no está determinada por ningún número racional. Para cada cortadura de ese tipo «creamos un nuevo miembro irracional α que está completamente definido por la cortadura; diremos que el número α corresponde a esa cortadura, o que produce la cortadura». Así, a cada cortadura le corresponde uno y un solo número, racional o irracional.

El lenguaje de Dedekind al introducir los números irracionales deja un poco que desear. Introduce el irracional α como correspondiente a la cortadura y definido por ella, pero no está demasiado claro de dónde viene α. Debería decir que el número irracional α

no es otra cosa que la cortadura misma. De hecho, Heinrich Weber le dijo esto a Dedekind, y este le replicó en una carta de 1888 que el número irracional α no es la cortadura misma, sino algo distinto, que corresponde a la cortadura y que la produce. De modo parecido, aunque los números racionales generan cortaduras, *no* son lo mismo que ellas. Y añadía que tenemos el poder mental de crear tales conceptos.

A continuación define cuándo una cortadura (A_1,A_2) es menor o mayor que otra cortadura (B_1,B_2). Después de haber definido la desigualdad, señala que los números reales poseen tres propiedades demostrables: 1) si $\alpha > \beta$ y $\beta > \gamma$, entonces $\alpha > \gamma$. 2) Si α y γ son dos números reales diferentes, entonces hay una cantidad infinita de números diferentes entre α y γ. 3) Si α es cualquier número real, entonces los números reales pueden dividirse en dos clases A_1 y A_2, cada una de las cuales contiene una infinita de elementos, y cada elemento de A_1 es menor que α y cada elemento de A_2 es mayor que α. El número α mismo puede ser asignado a cualquiera de las dos clases. El conjunto de los números reales posee ahora *continuidad*, que Dedekind expresa así: si se divide el conjunto de todos los números reales en dos clases A_1 y A_2 tales que cada elemento de A_1 es menor que todos los elementos de A_2, entonces existe un y sólo un número α que produce esa división.

Define después las operaciones con números reales. La suma de las cortaduras (A_1,A_2) y (B_1,B_2) se define así: si c es cualquier número racional, lo pondremos en la clase C_1 si hay un número a_1 en A_1 y un número b_1 en B_1 tales que $a_1 + b_1 \geqslant c$. Pondremos todos los demás números racionales en la clase C_2. Ese par de clases C_1 y C_2 constituyen una cortadura (C_1,C_2), puesto que cada elemento de C_1 es menor que cada elemento de C_2. La cortadura (C_1,C_2) es entonces la suma de (A_1,A_2) y (B_1,B_2). Las otras operaciones, dice, se definen análogamente. Ahora ya puede establecer propiedades de la suma y la multiplicación, como la propiedad asociativa y conmutativa. Aunque la teoría de Dedekind de los números irracionales, con pequeñas modificaciones como la señalada más arriba, es lógicamente satisfactoria, Cantor la criticó porque las cortaduras no aparecen de manera natural en análisis.

Hay otros enfoques de la teoría de los números irracionales además de los ya mencionados o descritos. Por ejemplo, en 1696 Wallis había identificado números racionales y números decimales periódicos. Otto Stolz (1842-1905), en su *Vorlesungen über allgemeine*

Arithmetik [14] mostró que cada número irracional puede representarse como un decimal no periódico, y esto puede utilizarse como propiedad definitoria.

Queda claro de estos varios enfoques que la definición lógica del número irracional es bastante refinada. Desde el punto de vista lógico un número irracional no es simplemente un solo símbolo o un par de símbolos, tal como una razón entre dos enteros, sino una colección infinita, como una sucesión fundamental de Cantor o una cortadura de Dedekind. El número irracional, definido lógicamente, es un monstruo intelectual, y podemos comprender por qué los griegos y tras ellos tantas generaciones de matemáticos los encontraron tan difíciles de manejar.

Los avances en matemáticas no son recibidos con universal aprobación. Hermann Hankel, creador él mismo de una teoría lógica de los números racionales, objetó a las teorías de los números irracionales:[15] «Cualquier intento de tratar los números irracionales formalmente y sin el concepto de magnitud (geométrica) conduce necesariamente a los artificios más abstrusos e incómodos, e incluso si son llevados a cabo con completo rigor, sobre lo que tenemos perfecto derecho a dudar, no tienen mucho valor científico.»

4. La teoría de los números racionales

La siguiente etapa en la fundamentación del sistema numérico fue la definición y deducción de las propiedades de los números racionales. Como ya hemos señalado, algún esfuerzo en esa dirección precedió al trabajo sobre los números irracionales. La mayoría de los que escribieron sobre números racionales supusieron que la naturaleza y propiedades de los enteros ordinarios eran conocidas, y que el problema era establecer lógicamente los números negativos y las fracciones.

El primer intento fue realizado por Martin Ohm (1792-1872), profesor en Berlín y hermano del físico, en su *Versuch eines vollkommen consequenten Systems der Mathematik* (Estudio de un sistema completo y consistente de las Matemáticas, 1822). Más tarde,

---

[14] 1886, 1, 109-119.

[15] *Theorie der complexen Zahlensystem*, 1867, 46-47.

Weierstrass, en sus lecciones de los años sesenta, presentaba los números racionales a partir de los naturales introduciendo los racionales positivos como pares de números naturales, los enteros negativos como otro tipo de pares de números naturales, y los racionales negativos como pares de enteros negativos. Esta idea fue utilizada independientemente por Peano, y la presentaremos más tarde con más detalle en relación con la obra de este último. Weierstrass no sintió la necesidad de clarificar la lógica de los números naturales. De hecho, su teoría de los números racionales no estaba exenta de dificultades. No obstante, en sus lecciones desde 1859 en adelante, afirmó correctamente que una vez admitidos los números naturales, ya no había necesidad de más axiomas para construir los números reales.

El problema clave para la construcción del sistema de los números racionales consistía en fundamentar los enteros ordinarios por algún procedimiento y en establecer sus propiedades. Entre los que trabajaron sobre la teoría de los enteros, algunos creían que eran tan fundamentales que no podía hacerse ningún análisis lógico de ellos. Esta posición fue adoptada por Kronecker, motivado por consideraciones filosóficas en las que después entraremos más profundamente. Kronecker también deseaba aritmetizar el análisis, esto es, fundamentar el análisis sobre los enteros; pero pensaba que no se podía ir más allá del reconocimiento de su existencia. El hombre los poseería mediante alguna especie de intuición fundamental. «Dios hizo los enteros», decía, «el resto es obra del hombre».

Dedekind ofreció una teoría de los enteros en su *Was sind und was sollen die Zahlen.*[16] Aunque fue publicada en 1888, la obra data de los años que van de 1872 a 1878. Utilizó en ella ideas de la teoría de conjuntos que Cantor había expuesto ya para esa época, y que iban a adquirir gran importancia. Sin embargo, el enfoque de Dedekind era tan complicado que se le prestó poca atención.

La aproximación a los enteros que mejor se adaptaba a las proclividades axiomatizadoras de finales del siglo XIX consistía en introducirlos mediante un conjunto de axiomas. Desconocedor en ese momento de los resultados obtenidos por Dedekind en la obra mencionada más arriba, Giuseppe Peano (1858-1932) también lo hizo en su *Arithmetices Principia Nova Methodo Exposita* (1889).[17] Como

---

[16] La naturaleza y significado de los números, en *Werke*, 3, 335-391.

[17] *Opere scelte*, 2, 20-56, y *Rivista di Matematica*, 1, 1891, 87-102, 256-257, en *Opere scelte*, 3, 80-109.

el enfoque de Peano es muy frecuentemente utilizado, lo expondremos aquí.

Buscando la precisión del razonamiento, utilizaba grandes dosis de simbolismo. Así, ε significa pertenecer a; $\supset$ significa implica; N_0 denota la clase de los números naturales, y $a+$ el número natural que viene a continuación de a. Peano utilizó este simbolismo en su presentación de todas las matemáticas, en particular en su *Formulario mathematico* (5 vols., 1895-1908), y también en sus clases, lo que provocó una revuelta de los estudiantes. Trató entonces de satisfacerles aprobándoles a todos, pero esto no funcionó, y fue obligado a renunciar a su puesto de profesor en la academia militar, permaneciendo a partir de entonces en la Universidad de Turín.

Aunque la obra de Peano influyó en el desarrollo posterior de la lógica simbólica y en el intento de Frege y Russell de asentar las matemáticas sobre la lógica, su obra debe distinguirse de la de estos últimos. Peano *no* quería fundamentar las matemáticas en la lógica, que para él sólo era una disciplina auxiliar de la matemática.

Peano partía de los conceptos no definidos (vid. cap. 42, sec. 2) de «conjunto», «número natural», «sucesor» y «pertenecer a». Sus cinco axiomas para los números naturales son:

1) 1 es un número natural.
2) 1 no es el sucesor de ningún otro número natural.
3) Cada número natural a tiene un sucesor.
4) Si los sucesores de a y b son iguales lo mismo pasa con a y b.
5) Si un conjunto S de números naturales contiene a 1, y cuando contiene a algún número natural a también contiene al sucesor de a, entonces S contiene todos los números naturales.

Este último axioma es el de la inducción matemática.

Peano también adoptó los axiomas de reflexividad, simetría y transitividad para la igualdad. Esto es, $a = a$; si $a = b$, entonces $b = a$; y si $a = b$ y $b = c$, entonces $a = c$. Definió la adición estableciendo que para cada par de números naturales a y b hay una suma única tal que

$$a + 1 = a+$$
$$a + (b+) = (a + b)+.$$

De modo parecido, la multiplicación quedaba definida estableciendo

que para cada par de números naturales a y b hay un producto único tal que

$$a \cdot 1 = a$$
$$a \cdot (b+) = (a \cdot b) + a.$$

A continuación exponía todas las propiedades acostumbradas de los números naturales.

A partir de los números naturales y sus propiedades es ya muy simple definir y establecer las propiedades de los números enteros negativos y de los números racionales. Se pueden definir primero los enteros positivos y negativos como una nueva clase de números, cada uno de ellos como un par ordenado de números naturales. Así (a,b), donde a y b son números naturales, es un entero. El significado intuitivo de (a,b) es $a - b$. Cuando $a > b$, el par representa un entero positivo ordinario, y cuando $a < b$, un entero negativo. Las definiciones adecuadas de las operaciones de adición y multiplicación conducen a las propiedades habituales de los enteros positivos y negativos.

Dados los enteros, se introducen los números racionales como pares ordenados de enteros. Así, si A y B son enteros, el par ordenado (A,B) es un número racional. Intuitivamente (A,B) es A/B. Una vez más, las definiciones adecuadas de adición y multiplicación de esos pares conducen a las propiedades usuales de los números racionales.

Así, una vez se hubo alcanzado el enfoque lógico de los números naturales, el problema de construir una fundamentación para el sistema numérico real quedaba resuelto. Como ya hemos señalado, las personas que trabajaron en la teoría de los irracionales suponían generalmente que los números racionales eran tan perfectamente conocidos que podían tomarlos como punto de partida, o en todo caso que bastaría con un pequeño esfuerzo para clarificarlos. Después de que Hamilton hubiera basado los números complejos en los reales, y después de que los irracionales hubieran sido definidos en función de los racionales, se creó finalmente la lógica de esta última clase de números. El orden histórico fue esencialmente el inverso del orden lógico requerido para la construcción del sistema de los números complejos.

5. Otros enfoques del sistema de los números reales

La esencia de los enfoques de la fundamentación lógica del sistema de los números reales descritos hasta el momento reside en la obtención de los enteros y sus propiedades por algún procedimiento para luego obtener, a partir de ellos, las fracciones y más tarde los números irracionales. La base lógica de este enfoque consiste en ciertas series de asertos únicamente acerca de los números naturales; por ejemplo, los axiomas de Peano. Para todos los demás números se ofrece un proceso de construcción. Hilbert llamó a ese enfoque «método genético» (puede que no conociera en ese momento los axiomas de Peano, pero sí otros similares para los números naturales). Reconociendo que el método genético puede tener valor pedagógico o heurístico, consideraba sin embargo como más seguro desde el punto de vista de la lógica la aplicación del método axiomático al sistema completo de los números reales. Antes de examinar sus razones, echemos un vistazo a sus axiomas [18].

Introduce el término no definido «número», denotado por a, b, c, ..., y plantea a continuación los siguientes axiomas:

I. Axiomas de conexión

I_1. A partir del número a y el número b se obtiene por adición un determinado número c; simbólicamente

$$a + b = c \qquad o \qquad c = a + b.$$

I_2. Si a y b son números dados, existe uno y sólo un número x y existe también un y sólo un número y tales que

$$a + x = b \qquad \text{y} \qquad y + a = b.$$

I_3. Hay un número determinado, denotado por 0, tal que para cualquier a

$$a + 0 = a \qquad \text{y} \qquad 0 + a = a.$$

---

[18] *Jahres. der Deut. Math.-Verein.*, 8, 1899, 180-184; este artículo no está en las *Gesammelte Abhandlungen* de Hilbert. Aparece en sus *Grundlagen der Geometrie*, 7.ª edición, Apéndice 6.

I_4. A partir del número a y el número b se obtiene por otro método, la multiplicación, un número determinado c; simbólicamente

$$ab = c \qquad \text{o} \qquad c = ab.$$

I_5. Dados dos números arbitrarios, a y b, si a no es 0, existe uno y sólo un número x, y también uno y sólo un número y, tales que

$$ax = b \qquad \text{a} \qquad ya = b.$$

I_6. Existe un número determinado, denotado por 1, tal que para cada a tenemos

$$a \cdot 1 = a \qquad \text{y} \qquad l \cdot a = a.$$

II. Axiomas de cálculo

II_1. $a + (b + c) = (a + b) + c$.
II_2. $a + b = b + a$.
II_3. $a(bc) = (ab)c$.
II_4. $a(b + c) = ab + ac$.
II_5. $(a + b)c = ac + bc$.
II_6. $ab = ba$.

III. Axiomas de orden

III_1. Si a y b son dos números diferentes, entonces uno de ellos es siempre mayor que el otro; este último se dice que es más pequeño que el primero; simbólicamente

$$a > b \qquad \text{y} \qquad b < a.$$

III_2. Si $a > b$ y $b > c$, entonces $a > c$.
III_3. Si $a > b$, entonces siempre es cierto que

$$a + c > b + c \qquad \text{y} \qquad c + a > c + b.$$

III_4. Si $a > b$ y $c > 0$, entonces $ac > bc$ y $ca > cb$.

IV. Axiomas de continuidad

IV_1. (Axioma de Arquímedes). Si $a > 0$ y $b > 0$ son dos núme-

ros arbitrarios, entonces es posible siempre sumar *a* consigo mismo un número suficiente de veces para tener

$$a + a + \ldots + a > b.$$

IV_2. (Axioma de completitud). No es posible añadir al sistema de los números ninguna colección de cosas de manera que la colección resultante siga satisfaciendo los axiomas precedentes; dicho en pocas palabras, los números forman una colección de objetos que no puede ampliarse sin que deje de cumplirse alguno de los axiomas precedentes.

Hilbert señala que esos axiomas no son independientes; se pueden deducir algunos de ellos de los otros. Afirma a continuación que la objeción contra la existencia de conjuntos infinitos (vid. sec. 6) no es válida para esta concepción de los números reales, ya que, dice, no tenemos que pensar en la colección de todas las leyes posibles según las que se pueden formar los elementos de una sucesión fundamental (sucesiones de Cantor de números racionales), sino únicamente en un sistema cerrado de axiomas y conclusiones que se pueden deducir de ellos tras un número finito de pasos lógicos. Señala desde luego que es necesario probar la consistencia de ese conjunto de axiomas, pero que desde el momento en que eso esté hecho, los objetos que los axiomas definen, es decir, los números reales, existirán desde el punto de vista matemático. Hilbert no era consciente en ese momento de la dificultad de probar la consistencia de los axiomas que había propuesto para los números reales.

A la afirmación de Hilbert de que el método axiomático es superior al genético, Bertrand Russell objetaba que ofrece la misma ventaja que el robo sobre el trabajo honrado: se arroga de inmediato todo lo que se puede construir por argumentos deductivos a partir de un conjunto mucho menor de axiomas.

Como en el caso de casi todos los avances significativos en matemáticas, la creación de la teoría de los números reales suscitó cierta oposición. Du Bois-Reymond, a quien ya hemos citado como oponente a la aritmetización del análisis, escribió en su *Théorie générale des fonctions* de 1887:[19]

Sin duda, con ayuda de los llamados axiomas, a partir de convenios, con

---

[19] Pág. 62 de la edición francesa de *Die allgemeine Funktionentheorie*, 1882.

proposiciones filosóficas construidas *ad hoc*, extendiendo ininteligiblemente conceptos originalmente claros, se puede construir un sistema aritmético que se parece en todos los aspectos al que se obtiene a partir del concepto de magnitud, para aislar así la matemática computacional, por decirlo de algún modo, mediante un cordón sanitario de dogmas y definiciones defensivas... Pero de esa forma se podrían inventar también otros sistemas aritméticos. La aritmética ordinaria no es otra que la que corresponde al concepto de magnitud lineal.

A pesar de ataques como éste, a los matemáticos les pareció que la culminación de los trabajos sobre los números reales resolvía todos los problemas lógicos que se habían presentado en el asunto. La aritmética, el álgebra y el análisis constituían en conjunto la mayor parte de las matemáticas, y esa parte quedaba ahora firmemente asentada.

6. El concepto de conjunto infinito

La rigorización del análisis había revelado la necesidad de comprender la estructura de los conjuntos de números reales. Para abordar este problema, Cantor había introducido ya (cap. 40, sec. 6) algunas nociones acerca de conjuntos infinitos de puntos, en particular los conjuntos de primera especie. Decidió que el estudio de los conjuntos infinitos era tan importante que debía dedicarse a examinar los conjuntos infinitos como tales, lo que le capacitaría, según creía, para distinguir claramente los diferentes conjuntos infinitos de discontinuidades.

La dificultad crucial en la teoría de conjuntos es el concepto mismo de conjunto infinito. Tales conjuntos habían suscitado naturalmente la atención de los matemáticos y los filósofos desde los tiempos de la antigua Grecia en adelante, y su naturaleza misma, así como sus propiedades aparentemente contradictorias, habían impedido cualquier progreso en su comprensión. Las paradojas de Zenón habían sido quizá la primera indicación de las dificultades. Ni la divisibilidad infinita de la recta ni la concepción de esta misma como conjunto infinito de puntos discretos parecían permitir conclusiones razonables acerca del movimiento. Aristóteles consideró conjuntos infinitos, como el de los números naturales, y negó la existencia de un conjunto infinito de objetos como entidad fija. Para él, los conjuntos sólo podían ser potencialmente infinitos (vid. cap. 3, sec. 10).

Proclo, el comentador de Euclides, señaló que como el diámetro de un círculo lo divide en dos mitades y hay un número infinito de diámetros, tendría que haber dos veces ese número de mitades. Esto les parece a muchos una contradicción, dice Proclo, y la resuelve diciendo que no se puede hablar de un infinito actual de diámetros o de mitades en un círculo, sino sólo de un número cada vez más grande de diámetros o de mitades. Con otras palabras, Proclo aceptaba el concepto aristotélico de un infinito potencial pero no actual, lo que evita el problema de una infinidad doble que iguala a una infinidad.

A lo largo de la Edad Media, los filósofos se situaron de un lado u otro con respecto a la cuestión de si puede haber una colección infinita actual de objetos. Se señaló que los puntos de dos circunferencias concéntricas podían ponerse entre sí en correspondencia uno a uno asociando los puntos del mismo radio. Y sin embargo una de las circunferencias es más larga que la otra.

Galileo se enfrentó a los conjuntos infinitos rechazándolos porque no podían ser sometidos a la razón. En su *Dos Nuevas Ciencias* (pp. 18-40 de la traducción inglesa) señala que los puntos de dos segmentos de distintas longitudes *AB* y *CD* pueden ponerse entre sí en correspondencia uno a uno (fig. 41.1) y por tanto contienen presumiblemente el mismo número de puntos. También señala que los números naturales pueden ponerse en correspondencia biunívoca con sus cuadrados asignando simplemente a cada número sus cuadrados; pero esto lleva a diferentes «cantidades» de infinito, lo que Galileo considera imposible. Todas las cantidades infinitas son la misma y no pueden compararse.

Gauss, en su carta a Schumacher del 12 de julio de 1831,[20] dice:

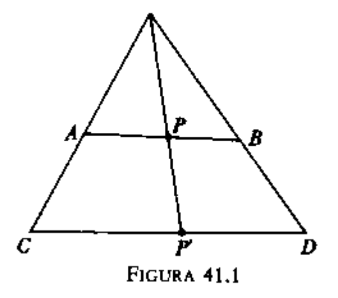


FIGURA 41.1

[20] *Werke*, 8, 216.

«Protesto contra el uso de una cantidad infinita como entidad actual; eso no está permitido nunca en matemáticas. El infinito es solamente una manera de hablar, en la que uno está realmente hablando de límites a los que ciertas razones pueden aproximarse tanto como se quiera, mientras que otras crecen ilimitadamente.» Cauchy, como otros antes que él, negaba la existencia de conjuntos infinitos porque el hecho de que una parte de ellos pudiera ponerse en correspondencia biunívoca con el todo le parecía contradictorio.

La polémica sobre los diferentes problemas concernientes a los conjuntos era interminable y recogía argumentos metafísicos e incluso teológicos. La actitud de la mayoría de los matemáticos hacia este problema consistía en ignorar lo que no podían resolver. En general, también evitaban el reconocimiento explícito de conjuntos actualmente infinitos, aunque utilizaban series infinitas, por ejemplo, y el sistema de números reales. Hablaban de los puntos de una recta evitando decir que ésta está formada por un número infinito de puntos. Este escamoteo de los problemas incómodos era hipócrita, pero suficiente para construir el análisis clásico. Sin embargo, cuando el siglo XIX se enfrentó al problema de introducir el rigor en análisis, ya no se pudieron evitar muchas cuestiones acerca de los conjuntos infinitos.

7. Los fundamentos de la teoría de conjuntos

Bolzano, en sus *Paradojas del Infinito* (1851), publicada tres años después de su muerte, fue el primero en dar pasos decisivos hacia una definitiva teoría de conjuntos. Defendió la existencia de conjuntos actualmente infinitos, y puso el acento en la noción de equivalencia de dos conjuntos, con la que aludía a lo que más tarde se llamaría correspondencia biunívoca entre los elementos de ambos conjuntos. Esta noción de equivalencia se podía aplicar tanto a conjuntos finitos como infinitos. Señaló que en el caso de conjuntos infinitos una parte o subconjunto podía ser equivalente al total, e insistió en que esto debía ser aceptado. Así, los números reales entre 0 y 5 pueden ser puestos en correspondencia biunívoca con los números reales entre 0 y 12 mediante la fórmula $y = 12x/5$, a pesar de que el segundo conjunto de números contiene al primero. A los conjuntos infinitos se les podrían atribuir números, y habría diferentes números transfinitos para los diferentes conjuntos infinitos,

si bien la asignación que Bolzano hacía de números transfinitos era incorrecta según la teoría posterior de Cantor.

La obra de Bolzano sobre el infinito era más filosófica que matemática, y no dejaba suficientemente clara la noción de lo que más tarde se llamaría potencia o número cardinal de un conjunto. También él encontró propiedades que la parecieron paradójicas, y las cita en su libro. Decidió que los números transfinitos no eran necesarios para fundamentar el cálculo, y en consecuencia no prosiguió más lejos la tarea.

El verdadero creador de la teoría de conjuntos fue Georg Cantor (1845-1918), que había nacido en Rusia en una familia judeo-danesa, pero se trasladó a Alemania con sus padres. Su progenitor le instó a estudiar ingeniería, y Cantor ingresó en la Universidad de Berlín en 1863 con esa intención. Allí fue influido por Weierstrass y se volcó en la matemática pura. Llegó a *Privatdozent* en Halle en 1869, y a profesor en 1879. Cuando tenía sólo veintinueve años publicó su primer artículo revolucionario sobre la teoría de conjuntos infinitos en el *Journal für Mathematik*. Aunque algunas de sus proposiciones fueron consideradas erróneas por los matemáticos más viejos, su total originalidad y brillo atrajeron la atención. Siguió publicando artículos sobre la teoría de conjuntos y los números transfinitos hasta 1897.

La obra de Cantor, que resolvía problemas muy antiguos y trastocaba mucho del pensamiento anterior, difícilmente podía recibir una aceptación inmediata. Sus ideas sobre los números transfinitos ordinales y cardinales suscitaron la hostilidad del poderoso Leopold Kronecker, que atacó brutalmente las ideas de Cantor durante más de una década. En cierto momento Cantor sufrió una crisis nerviosa, pero reemprendió el trabajo en 1887. Aunque Kronecker murió en 1891, sus ataques despertaron las sospechas de los matemáticos acerca de la obra de Cantor.

Su teoría de conjuntos aparece dispersa en muchos artículos, y no intentaremos por tanto indicar el lugar específico en el que aparece cada uno de sus conceptos y teoremas. Esos artículos aparecieron en los *Mathematische Annalen* y el *Journal für Mathematik* desde 1874 en adelante.[21] Por conjunto Cantor entiende una colección de objetos definidos y separados que puede ser concebida por la mente y sobre la que podemos decidir si un determinado objeto

---

[21] *Ges. Abh.*, 115-356.

pertenece o no a ella. Mantiene que los que defienden conjuntos sólo potencialmente infinitos se equivocan, y refuta los argumentos previos de matemáticos y filósofos contra los conjuntos actualmente infinitos. Para Cantor un conjunto es infinito si puede ponerse en correspondencia biunívoca con una parte de sí mismo. Algunos de sus conceptos, como el de punto límite de un conjunto, o los de conjunto derivado, conjunto de primera especie, etc., fueron definidos y utilizados en un artículo sobre series trigonométricas,[22] que ya hemos descrito en el capítulo precedente (sec. 6). Un conjunto es cerrado si contiene todos sus puntos límite; es abierto si todos sus puntos son interiores, esto es, si cada uno de sus puntos puede ser incluido en un intervalo que sólo contiene puntos del conjunto; es perfecto si cada uno de sus puntos es un punto límite y además es cerrado. También definió la unión e intersección de conjuntos. Aunque estaba ante todo interesado en conjuntos de puntos de una recta o conjuntos de números reales, extendió esos conceptos a los conjuntos de puntos de un espacio euclídeo n-dimensional.

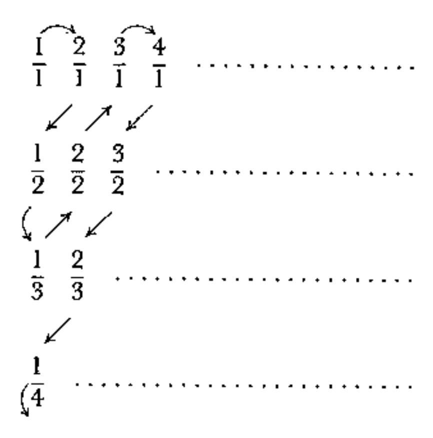
Trató después de distinguir los conjuntos infinitos según su «tamaño» y, como Bolzano, decidió que la correspondencia uno a uno debía ser el criterio básico. Dos conjuntos que pueden ponerse en correspondencia biunívoca son equivalentes, o tienen la misma potencia (más tarde, la palabra «potencia» sería reemplazada por «número cardinal»). Dos conjuntos pueden tener distinta potencia; si de dos conjuntos de objetos M y N, N puede ponerse en correspondencia biunívoca con un subconjunto de M, pero M no puede ponerse en correspondencia biunívoca con ningún subconjunto de N, la potencia de M es mayor que la de N.

Los conjuntos de números eran desde luego los más importantes, y por eso Cantor ilustra su concepto de equivalencia o potencia con tales conjuntos. Introduce el término «numerable» para cualquier conjunto que pueda ponerse en correspondencia biunívoca con los enteros positivos. Este es el conjunto infinito más pequeño. Cantor probó que el conjunto de los números racionales es numerable. Ofreció una demostración en 1874.[23] No obstante, su segunda demostración [24] es la más empleada, y es la que vemos a describir.

---

[22] *Math. Ann.*, 5, 1872, 122-132 = *Ges. Abh.*, 92-102.
[23] *Jour. für Math.*, 77, 1874, 258-262 = *Ges. Abh.*, 115-118.
[24] *Math. Ann.*, 46, 1895, 481-512 = *Ges. Abh.*, 283-356, en particular 294-295.

Dispone los números racionales así:



Puede observarse que todos los que hay en cualquier diagonal tienen la misma suma de numerador y denominador. Parte entonces de 1/1 y sigue las flechas asignando el número 1 a 1/1, 2 a 2/1, 3 a 1/2, 4 a 1/3, y así sucesivamente. Cada número racional será alcanzado en algún momento, y a cada uno se le asignará un entero finito. Luego el conjunto de más arriba de números racionales (en el que algunos aparecen muchas veces) está en correspondencia uno a uno con los enteros. Si se suprimen las repeticiones, el conjunto de los números racionales será todavía infinito y necesariamente numerable, ya que éste es el menor conjunto infinito.

Todavía más sorprendente es la demostración de Cantor en el artículo citado de 1874 de que el conjunto de todos los números algebraicos, esto es, el conjunto de todos los números que son soluciones de ecuaciones algebraicas

$$a_o x^n + a_1 x^{n-1} + \dots + a_n = 0,$$

donde los a_i son enteros, es también numerable.

Para probar esto, asigna a cada ecuación algebraica de grado n la altura N definida por

$$N = n - 1 + |a_o| + |a_1| + \dots + |a_n|,$$

donde los a_i son los coeficientes de la ecuación. La altura N es un

entero. A cada N corresponden solamente un número finito de ecuaciones algebraicas, y por tanto solamente una cantidad finita de números algebraicos, digamos $\varphi(N)$. Por ejemplo, $\varphi(1) = 1$; $\varphi(2) = 2$; $\varphi(3) = 4$. A partir de $N = 1$, numera los correspondientes números algebraicos desde 1 hasta n_1; los números algebraicos de altura 2 son numerados desde $n_1 + 1$ hasta n_2; y así sucesivamente. Como cada número algebraico será alcanzado en alguna etapa y se le asignará uno y sólo un entero, el conjunto de los número algebraicos es numerable.

En sus cartas a Dedekind en 1873, Cantor le planteó la cuestión de si el conjunto de los números reales podría ponerse en correspondencia uno a uno con los enteros, y algunas semanas más tarde le aseguró no era posible. Ofreció dos demostraciones. La primera (en el artículo de 1874 que ya hemos citado) era más complicada que la segunda,[25] que es la que más a menudo se expone en la actualidad. Tiene también la ventaja, como señaló Cantor, de ser independiente de consideraciones técnicas acerca de los números irracionales.

La segunda demostración de Cantor de que los números reales son no numerables (no enumerables) comienza suponiendo que se pueden enumerar los números reales entre 0 y 1. Escribamos cada uno de ellos en forma decimal, conviniendo en que un número tal como 1/2 se escribirá 0,4999... Si esos números reales son numerables, podremos asignar cada uno de ellos a un entero n, así:

$$\begin{array}{ccl} 1 & \leftrightarrow & 0,\ a_{11}\ a_{12}\ a_{13}\ \ldots \\ 2 & \leftrightarrow & 0,\ a_{21}\ a_{22}\ a_{23}\ \ldots \\ 3 & \leftrightarrow & 0,\ a_{31}\ a_{32}\ a_{33}\ \ldots \end{array}$$

Definamos ahora un número real entre 0 y 1 así: sea $b = 0.\ b_1\ b_2\ b_3, \ldots,$ donde $b_k = 9$ si $a_{kk} = 1$ y $b_k = 1$ si $a_{kk} \neq 1$. Ese número real difiere de cada uno de los escritos en la correspondencia de más arriba. Sin embargo, se suponía que esa lista contenía todos los números reales entre 0 y 1. Hemos llegado así a una contradicción.

Como los números reales no son numerables y los números algebraicos son numerables, tiene que haber irracionales trascendentes. Esta es la demostración de existencia no constructiva de Cantor, que podríamos comparar con la construcción efectiva de irracionales trascendentes de Liouville (sec. 2).

---

[25] *Jahres. der Deut. Math.-Verein.*, 1, 1890/1891, 75-78 = *Ges. Abh.*, 278-281.

En 1874 Cantor se ocupó de la equivalencia entre los puntos de una recta y los de $\mathbb{R}^n$ (espacio n-dimensional) tratando de probar que una correspondencia biunívoca entre esos dos conjuntos era imposible. Tres años más tarde comprobó que sí existe tal correspondencia, escribiendo a Dedekind: «Lo veo, pero no lo creo.»[26]

La idea utilizada para establecer esa correspondencia uno a uno [27] puede generalizarse fácilmente a partir de una correspondencia entre los puntos del cuadrado unidad y los del segmento (0,1). Sean (x,y) un punto del cuadrado unidad y z un punto del intervalo unidad. Representemos x e y en forma decimal, reemplazando los ceros de un decimal finito por una sucesión infinita de nueves. A continuación separemos en x e y grupos de cifras decimales que terminen con la primera cifra no nula de la sucesión. Por ejemplo,

$$x = 0,\ 3\ 002\ 03\ 04\ 6\ \ldots$$
$$y = 0,\ 01\ 6\ 07\ 8\ 09\ \ldots$$

Formemos entonces

$$z = 0,\ 3\ 01\ 002\ 6\ 03\ 07\ 04\ 8\ 6\ 09\ \ldots$$

eligiendo como grupos en la expresión decimal de z el primer grupo de x, a continuación el primero grupo de y, y así sucesivamente. Si las coordenadas x o y de dos puntos del cuadrado difieren en alguna cifra decimal, los z correspondientes serán diferentes. Luego para cada (x,y) hay un único z. Recíprocamente, dado un z, se pueden separar en su expresión decimal los grupos descritos anteriormente, formando x e y, de manera que para cada z hay un único (x,y). La correspondencia biunívoca que acabamos de establecer no es continua; *grosso modo*, eso significa que a puntos z próximos no les corresponden necesariamente puntos (x,y) próximos, ni a la inversa.

Du Bois-Reymond objetó a esta demostración: [28] «Repugna al sentido común. De hecho, se trata simplemente de la conclusión de un tipo de razonamiento que permite la intervención de ficciones ideales, a las que se hace jugar el papel de cantidades genuinas aun-

---

[26] *Briefwechsel Cantor-Dedekind*, 34.
[27] *Jour. für Math.*, 84, 1878, 242-258 = *Ges. Abh.*, 119-133.
[28] Pág. 167 de la edición francesa (1887) de su *Die allgemeine Funktionentheorie*.

que no sean siquiera límites de representaciones de cantidades. Ahí es donde reside la paradoja.»

8. Cardinales y ordinales transfinitos

Habiendo demostrado la existencia de conjuntos con la misma potencia y con diferentes potencias, Cantor prosiguió con este concepto de potencia de un conjunto e introdujo una teoría de números cardinales y ordinales en la que los elementos destacados son los cardinales y ordinales transfinitos. Cantor desarrolló este trabajo en una serie de artículos en los *Mathematische Annalen* desde 1879 hasta 1884, bajo el título común «Uber unendliche lineare Punktmannichfaltigkeiten» (sobre agregados lineales infinitos de puntos). Después escribió dos artículos definitivos en 1895 y 1897 en la misma revista.[29]

En el quinto artículo sobre agregados lineales [30], Cantor comienza con la observación:

> La descripción de mis investigaciones en la teoría de agregados ha alcanzado un estado en el que su prolongación depende de una generalización de los enteros reales positivos más allá de sus límites actuales; una generalización en una dirección en la que, por lo que yo sé, nadie se ha aventurado todavía.
>
> Dependo de esa generalización del concepto de número hasta tal punto que sin ella no podría dar ni siquiera pequeños pasos adelante en la teoría de conjuntos. Espero que esta situación justifique o, si es necesario, excuse la introducción de ideas aparentemente tan extrañas en mis argumentaciones. De hecho, el objetivo consiste en generalizar o extender la serie de los enteros reales más allá del infinito. Por atrevido que esto pueda parecer, tengo no sólo la esperanza, sino la firme convicción de que a su debido tiempo esta generalización será reconocida como un paso bastante simple, apropiado y natural. Aun así, soy muy consciente de que adoptando tal procedimiento me sitúo a contracorriente con respecto a las opiniones generales sobre el infinito en matemáticas y sobre la naturaleza de los números.

---

[29] *Math. Ann.*, 46, 1895, 481-512, y 49, 1897, 207-246 = *Ges. Abh.*, 282-351; se puede encontrar una traducción al inglés de estos dos artículos en Georg Cantor, *Contributions to the Founding of a Theory of Transfinite Numbers,* Dover (reimpresión), sin fecha.

[30] 1883, *Ges. Abh.*, 165.

Señala que su teoría de los números infinitos o transfinitos es distinta del concepto de infinidad bajo el que se entiende una variable que se hace infinitamente pequeña o infinitamente grande. Dos conjuntos que están en correspondencia uno a uno tienen la misma potencia, o el mismo número cardinal. Para conjuntos finitos el número cardinal es el número usual de objetos del conjunto. Para conjuntos infinitos se introducen nuevos números cardinales. El denotó por $\aleph_0$ al del conjunto de los números enteros. Como los números reales no se pueden poner en correspondencia biunívoca con los números enteros, el conjunto de los números reales debe tener otro cardinal distinto, que se denota por *c*, la inicial de continuo. Como en el caso del concepto de potencia, si dos conjuntos *M* y *N* son tales que *N* puede ponerse en correspondencia biunívoca con un subconjunto de *M*, pero *M* no puede ponerse en correspondencia biunívoca con ningún subconjunto de *N*, el número cardinal de *M* es mayor que el de *N*. Así, $c > \aleph_0$.

Para obtener un número cardinal mayor que otro dado,[31] se considera cualquier conjunto *M* que represente a ese cardinal dado. Sea entonces *N* el conjunto de todos los subconjuntos de *M*. Entre esos subconjuntos están los elementos individuales de *M*, los pares de elementos de *M*, y así sucesivamente. Ahora bien, es evidentemente posible establecer una correspondencia biunívoca entre *M* y un subconjunto de *N*, el constituido por los elementos individuales de *M* (considerados como subconjuntos de *M*, y como elementos de *N*). Por otra parte, no es posible establecer una correspondencia uno a uno entre *M* y *N*. Supongamos efectivamente que sí la hubiera; sea *m* cualquier elemento de *M* y consideremos aquellos *m* tales que el elemento de *N* asociado por la supuesta correspondencia *no* contiene a ese *m*. Sea η el conjunto de tales *m*, que naturalmente es un elemento de *N*. Cantor afirma entonces que η no está incluido en la supuesta correspondencia uno a uno, ya que si η correspondiera a algún *m* de *M*, y η contuviera a *m*, tendríamos una contradicción con la definición de η. Pero si *m* no pertenece a η, que es el elemento de *N* que supuestamente le corresponde, como η era por definición el conjunto de todos los *m* no contenidos en los correspondientes elementos de *N*, tendría que contener a ese *m*. Así, la suposición de que hay una correspondencia biunívoca entre los elementos de *M* y los de *N*, que son los subconjuntos de *M*, conduce a una contradic-

---

[31] Cantor, *Ges. Abh.*, 278-280.

ción. Luego el número cardinal del conjunto que consiste en todos los subconjuntos de un conjunto dado es mayor que el cardinal de este último.

Cantor definió la suma de dos números cardinales como el cardinal de la unión de dos conjuntos (disjuntos) que representen a los sumandos. También definió el producto de dos cardinales cualesquiera: dados dos cardinales α y β, se toman un representante M de α y un representante N de β, y se forman los pares de elementos (m,n) donde m es un elemento de M y n de N. El producto de α y β es entonces el cardinal del conjunto de todos esos posibles pares.

Se definen también potencias de números cardinales. Si tenemos un conjunto M de m objetos y otro N de n objetos, Cantor define el conjunto m de las permutaciones de m objetos tomando n cada vez y permitiendo repeticiones de los m objetos iniciales. Así, por ejemplo, si $m = 3$ y $n = 2$, con los objetos m_1, m_2, m_3, las permutaciones que resultan son

$$\begin{matrix} m_1m_1 & m_2m_2 & m_3m_1 \\ m_1m_2 & m_2m_1 & m_3m_2 \\ m_1m_3 & m_2m_3 & m_3m_3 \end{matrix}$$

Cantor define α^β como el cardinal de este conjunto de permutaciones, siendo α el cardinal de M y β el de N. A continuación demuestra que $2^{\aleph_0} = c$, cardinal del continuo.

Cantor llama la atención sobre el hecho de que su teoría de números cardinales se aplica en particular a los cardinales finitos, proporcionando así «el fundamento más corto y más riguroso para la teoría de números finitos».

El siguiente concepto es el de número ordinal. Había encontrado ya la necesidad de tal concepto al introducir los sucesivos conjuntos derivados de un conjunto dado de puntos. Ahora los introduce abstractamente. Un conjunto está totalmente ordenado si para cada dos elementos uno de ellos precede al otro, de manera que dados m_1 y m_2, o bien m_1 precede a m_2, o m_2 precede a m_1: la notación es $m_1 < m_2$ o $m_2 < m_1$. Además, si $m_1 < m_2$ y $m_2 < m_3$, ese orden total también implica que $m_1 < m_3$; esto es, la relación de orden es transitiva. El número ordinal de un conjunto ordenado M es el tipo de orden del orden definido en el conjunto. Dos conjuntos ordenados son semejantes si hay una correspondencia biunívoca entre ellos

y si, cuando a m_1 le corresponde n_1 y a m_2 le corresponde n_2, y $m_1 < m_2$, entonces $n_1 < n_2$. Dos conjuntos ordenados semejantes tienen el mismo tipo o número ordinal. Como ejemplos de conjuntos ordenados podemos utilizar cualquier conjunto finito de números en cualquier orden dado. Para un conjunto finito, sea cual sea el orden, el número ordinal es el mismo y se puede tomar como símbolo para él el número cardinal del conjunto. El número ordinal del conjunto de los enteros positivos en su orden natural se denota por ω. El mismo conjunto de los enteros positivos, pero en orden decreciente, esto es,

$$..., 4, 3, 2, 1$$

se denota por $^*\omega$. El conjunto de todos los enteros, positivos y negativos, más el cero, en su orden natural, tiene como número ordinal $^*\omega + \omega$.

Cantor define a continuación la suma y el producto de números ordinales. La suma es el número ordinal del conjunto ordenado que se obtiene posponiendo al primero el segundo de los conjuntos dados, manteniendo en cada uno de ellos el orden original. Así, el conjunto de los enteros positivos seguido de los primeros cinco enteros, esto es,

$$1, 2, 3, ..., 1, 2, 3, 4, 5$$

tiene como número ordinal $\omega + 5$. También se definen de manera bastante obvia la igualdad y desigualdad de números ordinales.

Y ahora es cuando introduce el conjunto de todos los ordinales transfinitos, en parte por su propia importancia, y en parte para poder definir con precisión números cardinales transfinitos más elevados. Para introducir esos nuevos ordinales restringe los conjuntos totalmente ordenados a conjuntos bien ordenados.[32] Un conjunto está bien ordenado si tiene un primer elemento en la ordenación y también lo tiene cada uno de sus subconjuntos. Hay una jerarquía de números ordinales y números cardinales. En la primera clase, denotada por Z_1, están los ordinales finitos

$$1, 2, 3, ...$$

[32] *Math. Ann.*, 21, 1883, 545-586 = *Ges. Abh.*, 165-204.

En la segunda clase, denotada por Z_2, están los ordinales

$$\omega, \omega+1, \omega+2,...,2\omega, 2\omega+1,...,3\omega, 3\omega+1,...,\omega^2, \omega^3,...,\omega^\omega,...$$

Cada uno de esos ordinales es el ordinal de un conjunto cuyo cardinal es $\aleph_0$.

El conjunto de los ordinales que hay en Z_2 tiene a su vez un número cardinal. Ese conjunto no es numerable, y Cantor introduce para él un nuevo número cardinal $\aleph_1$, que muestra que es el primer cardinal posterior a $\aleph_0$.

Los ordinales de la tercera clase, denotada por Z_3, son

$$\Omega, \Omega + 1, \Omega+2,...,\Omega + \Omega, ...$$

Estos son números ordinales de conjuntos bien ordenados, cada uno de los cuales tiene $\aleph_1$ elementos. Sin embargo, el conjunto de ordinales Z_3 tiene más de $\aleph_1$ elementos, y Cantor denota su cardinal por $\aleph_2$. Esta jerarquía de ordinales y cardinales puede proseguirse indefinidamente.

Cantor había mostrado que dado cualquier conjunto siempre es posible crear otro, el conjunto de las partes del conjunto dado, cuyo cardinal es mayor que el de este último. Si el conjunto dado es $\aleph_0$, el número cardinal del conjunto de sus partes es $2^{\aleph_0}$. Cantor había demostrado también que $2^{\aleph_0} = c$, el cardinal del continuo. Por otra parte había introducido $\aleph_1$ mediante los números ordinales, y había probado que $\aleph_1$ es el primer cardinal que sigue a $\aleph_0$. Así pues, $\aleph_1 \leqslant c$, pero la cuestión era si $\aleph_1 = c$, cuestión conocida como hipótesis del continuo. A pesar de intentarlo esforzadamente, Cantor no pudo dar una respuesta. Hilbert incluyó esta cuestión en una lista de problemas sobresalientes presentada en el Congreso Internacional de Matemáticos de 1900 (vid. cap. 43, sec. 5 y cap. 51, sec. 8).

Para dos conjuntos en general M y N cabe la posibilidad de que M no pueda ponerse en correspondencia biunívoca con ningún subconjunto de N, sin que tampoco N se pueda poner en correspondencia biunívoca con ningún subconjunto de M. En tal caso, para los números cardinales α y β de M y N no se puede decir que $\alpha = \beta$, $\alpha < \beta$ o $\alpha > \beta$, esto es, esos dos cardinales no son comparables. Cantor fue capaz de probar que esta situación no puede darse si los conjuntos están bien ordenados. Parecía paradójico que pudiera haber conjuntos no bien ordenados cuyos cardinales no pu-

dieran compararse, pero tampoco pudo Cantor resolver este problema.

Ernst Zermelo (1871-1953) afrontó el problema de qué hacer para comparar los números cardinales de conjuntos que no estén bien ordenados. En 1904 [33] probó, ofreciendo en 1908 [34] una segunda demostración, que a todo conjunto se le puede dar una buena ordenación. En la demostración tuvo que utilizar lo que se conoce ahora como axioma de elección (axioma de Zermelo), según el cual, dada cualquier colección de conjuntos no vacíos disjuntos, es posible elegir en cada uno de ellos un elemento para formar un nuevo conjunto. El axioma de elección, el teorema sobre la buena ordenación, y el hecho de que dos conjuntos cualesquiera pueden compararse en cuanto a su tamaño (esto es, que si sus cardinales son α y β, o es $\alpha = \beta$, o $\alpha < \beta$, o bien $\alpha > \beta$), son principios equivalentes.

9. La situación de la teoría de conjuntos hacia 1900

La teoría de conjuntos de Cantor constituyó un audaz paso adelante en un dominio que, como ya hemos señalado, había sido explorado intermitentemente desde el tiempo de los griegos. Exigía una aplicación estricta de argumentos puramente racionales, y afirmaba la existencia de conjuntos infinitos de potencia tan elevada como se quiera, completamente fuera del alcance de la intuición humana. Habría sido extraño que tales ideas, mucho más revolucionarias que la mayoría de las introducidas anteriormente, no hubieran encontrado oposición. Las dudas en cuanto a la coherencia de este desarrollo se vieron reforzadas por ciertos interrogantes planteados por el mismo Cantor y por algunos otros. En sus cartas a Dedekind del 28 de julio y el 28 de agosto de 1899,[35] Cantor se preguntaba si el conjunto de todos los números cardinales podría ser realmente un conjunto, ya que entonces su cardinal sería mayor que cualquier otro. Pensó que la respuesta correcta debía ser la negativa, distinguiendo entre conjuntos consistentes e inconsistentes. Pero en 1897 Cesare Burali-Forti (1861-1931) señaló que la sucesión de *todos* los números ordinales, que está bien ordenada, debería tener como nú-

---

[33] *Math. Ann.*, 59, 1904, 514-516.
[34] *Math. Ann.*, 65, 1908, 107-128.
[35] *Ges. Abh.*, 445-448.

mero ordinal el mayor de todos los ordinales.[36] Pero entonces ese número ordinal sería mayor que *todos* los números ordinales (Cantor había apreciado ya esta dificultad en 1895). Estos y otros problemas no resueltos, llamados paradojas, empezaron a hacerse sentir a finales del siglo XIX.

La oposición a las ideas de Cantor crecía. Kronecker, como ya hemos dicho, se opuso casi desde el principio. Felix Klein tampoco simpatizaba en absoluto con ellas. Poincaré [37] observaba críticamente: «Pero sucede que hemos encontrado ciertas paradojas, ciertas contradicciones aparentes que habrían hecho las delicias de Zenón de Elea y de la escuela de Megara... Creo, por mi parte, y no soy el único, que el punto delicado está en la introducción de objetos que no pueden definirse completamente con un número finito de palabras.» Se refería a la teoría de conjuntos como un interesante «caso patológico», y predecía (en el mismo artículo) que «las generaciones posteriores considerarán las *Mengenlehre* (de Cantor) como una enfermedad de la que uno se ha curado». Hermann Weyl decía de la jerarquía de alefs de Cantor que era como una niebla en medio de la niebla.

No obstante, muchos matemáticos prominentes quedaron impresionados por la utilización que podía ya hacerse de la nueva teoría. En el primer Congreso Internacional de Matemáticos en Zurich (1897), Adolf Hurwitz y Hadamard señalaron importantes aplicaciones al análisis de la teoría de los números transfinitos. Pronto se descubrieron nuevas aplicaciones en la teoría de la medida (cap. 44) y en topología (cap. 50). Hilbert difundió las ideas de Cantor en Alemania, en 1926 [38] decía: «Nadie podrá expulsarnos del paraíso que Cantor creó para nosotros», y elogiaba la aritmética transfinita de Cantor como «el producto más impresionante del pensamiento matemático, una de las más bellas realizaciones de la actividad humana en el dominio de lo puramente inteligible».[39] Bertrand Russell

---

[36] *Rendiconti del Circolo Matematico di Palermo*, 11, 1897, 154-164 y 260.

[37] *Proceedings of the Fourth Internat. Cong. of Mathematicians*, Roma, 1908, 167-182. *Bull. des Sci. Math.* (2), 32, 1908, 168-190 = Hay un extracto de *Œuvres*, 5, 19-23.

[38] *Math. Ann.*, 95, 1926, 170 = *Grundlagen der Geometrie*, 7.ª edición, 1930, 274.

[39] *Math. Ann.*, 95, 1926, 167 = *Grundlagen der Geometrie*, 7.ª edición, 1930, 270. El artículo «Uber das Unendliche», del que hemos tomado las citas de más arriba, aparece también en francés en *Acta Math.*, 48, 1926, 91-122. No está incluido en las *Gesammelte Abhandlungen* de Hilbert.

describió la obra de Cantor como «la que probablemente puede enorgullecer más a nuestra época».

Bibliografía

Becker, Oskar: *Grundlagen der Mathematik in geschichtlicher Entwicklung*, Verlag Karl Alber, 1954, 217-316.

Boyer, Carl B.: *Historia de la matemática*, Madrid, Alianza, 1986.

Cantor, Georg: *Gesammelte Abhandlungen*, 1932, Georg Olms (reimpresión), 1962.

— *Contributions to the Founding of the Theory of Transfinite Numbers;* Dover (reimpresión), sin fecha. Contiene una traducción al inglés de los dos artículos clave de 1895 y 1897 y una muy útil introducción de P. E. B. Jourdain.

Cavailles, Jean: *Philosophie mathématique*, Hermann, 1962. Contiene también la correspondencia entre Cantor y Dedekind, traducida al francés.

Dedekind, R.: *Essays on the Theory of Numbers*, Dover (reimpresión), 1963. Contiene una traducción al inglés de los artículos de Dedekind «Stetigkeit und irrationale Zahlen» y «Was sind und was sollen die Zahlen». Ambos están también en sus *Werke*, 3, 314-334 y 335-391.

Fraenkel, Abraham A.: «Georg Cantor», *Jahres. der Deut. Math.-Verein.*, 39, 1930, 189-266. Revisión histórica de la obra de Cantor.

Helmholtz, Hermann von: *Counting and Measuring*, D. Van Nostrand, 1930. Traducción al inglés de la obra de HELMHOLTZ, *Zählen und Messen, Wissenchaftliche Abhandlungen*, 3, 356-391.

Manheim, Jerome H.: *The Genesis of Point Set Topology*, MacMillan, 1964, 76-110.

Meschkowski, Herbert: *Ways of Thought of Great Mathematicians*, Holden-Day, 1964, 91-104.

— *Evolution of Mathematical Thought*, Holden-Day, 1965, caps. 4-5.

— *Probleme des Unendlichen: Werk und Leben Georg Cantors*, F. Vieweg und Sohn, 1967.

Noether, E., y J. Cavaillés: *Briefwechsel Cantor-Dedekind*, Hermann, 1937.

Peano, G.: *Opere scelte*, 3 vols., Edizioni Cremonese, 1957-1959.

Schoenflies, Arthur M.: *Die Entwickelung der Mengenlehre und ihre Anwendungen*, en dos partes, B. G. Teubner, 1908 y 1913.

Smith, David Eugene: *A Source Book in Mathematics*, Dover (reimpresión), 1959, vol. 1, 35-45 y 99-106.

Stammler, Gerhard: *Der Zahlbegriff seit Gauss*, Georg Olms, 1965.

Capítulo 42
LOS FUNDAMENTOS DE LA GEOMETRIA

La geometría no es nada si no es rigurosa... Los métodos de Euclides son considerados casi universalmente como irreprochables desde el punto de vista del rigor.

H. J. S. SMITH (1873)

Ha sido costumbre defender a Euclides, cuando se le ataca como libro de texto por su verbosidad, su oscuridad o su pendatería, con el argumento de su excelencia lógica, que supuestamente proporcionaría un entrenamiento incomparable a las jóvenes capacidades de razonamiento. Esta suposición pierde consistencia, sin embargo, cuando se la examina más de cerca. Sus definiciones no siempre definen, sus axiomas no siempre son indemostrables, sus demostraciones requieren muchos axiomas de los que es enteramente inconsciente. Una prueba válida mantiene su poder demostrativo cuando no se dibuja ninguna figura, pero muchas de las demostraciones de Euclides no pasarían esa criba... El valor de su obra como obra maestra de la lógica se ha exagerado enormemente.

BERTRAND RUSSELL (1902)

1. Los defectos de Euclides

Las críticas a las definiciones y axiomas de Euclides (cap. 4, sec. 10) se remontan a sus primeros comentaristas conocidos, Pappus y Proclo. Cuando los europeos comenzaron a interesarse por Eu-

clides durante el Renacimiento, también apreciaron fallos. Jacques Peletier (1517-1582), en su *In Euclidis Elementa Geometrica Demonstrationum* (1557), criticó el uso que hace Euclides de la superposición para probar teoremas sobre congruencia. También el filósofo Arthur Schopenhauer manifestó en 1844 su sorpresa ante el hecho de que los matemáticos cuestionaran el postulado de las paralelas y no el axioma según el que las figuras que coinciden son iguales, argumentando que o bien las figuras coincidentes son automáticamente idénticas o iguales y entonces no se necesita ningún axioma, o bien la coincidencia es algo completamente empírico, que no pertenece a la pura intuición *(Anschauung)* sino a la experiencia sensorial externa. Además, el axioma presupone la movilidad de las figuras; pero lo que se puede mover en el espacio es materia, y queda por tanto fuera de la geometría. En el siglo XIX llegó a reconocerse generalmente que el método de superposición descansaba sobre axiomas no explicitados, o que debería reemplazarse por otro enfoque de la congruencia.

A algunos críticos no les gustaba como axioma la afirmación de que todos los ángulos rectos son iguales, y trataron de probarla, naturalmente sobre la base de los demás axiomas. Christophorus Clavius (1537-1612), uno de los editores de la obra de Euclides, señaló la ausencia de un axioma que garantizara la existencia de una cuarta proporcional para tres magnitudes dadas (cap. 4, sec. 5). Leibniz comentó acertadamente que Euclides se basaba en la intuición al afirmar (Libro I, Proposición 1) que dos circunferencias, cada una de las cuales pasa por el centro de la otra, tienen un punto en común. Con otras palabras, Euclides suponía que una circunferencia tiene una cierta estructura continua que le hace poseer un punto en el que la otra circunferencia la corta.

También Gauss señaló las imperfecciones de la presentación que Euclides hacía de la geometría. En una carta a Wolfgang Bolyai del 6 de marzo de 1832,[1] le hacía notar que hablar de la parte del plano limitada por un triángulo exige una fundamentación adecuada. Y decía también: «En un desarrollo completo, palabras tales como «entre» deben basarse en concepto claros, cosa que puede hacerse, pero que yo no he encontrado en ningún sitio.» Gauss realizó otras críticas sobre la definición de línea recta[2] y sobre la definición de

---

[1] *Werke*, 8, 222.
[2] *Werke*, 8, 196.

plano como una superficie en la que debe estar contenida la recta que une dos puntos cualesquiera del plano.[3]

Es bien sabido que pueden obtenerse muchas «demostraciones» de resultados falsos porque los axiomas de Euclides no indican dónde deben estar ciertos puntos con relación a otros. Está, por ejemplo, la «demostración» de que todo triángulo es isósceles: se construyen la bisectriz en *A* del triángulo *ABC* y la mediatriz del lado *BC* (fig. 42.1). Si esas dos rectas son paralelas, la bisectriz es perpendicular a *BC* y el triángulo es isósceles. Supongamos, por el contrario, que esas dos rectas se cortan, digamos en O, y vamos entonces a «probar» que el triángulo sigue siendo isósceles. Dibujemos las perpendiculares *OF* a *AB* y *OE* a *AC*.

Los triángulos señalados con I son congruentes, y $OF = OE$. Los triángulos señalados con III son también congruentes, y $OB = OC$. En consecuencia, los triángulos marcados con II son congruentes, y $FB = EC$. En los triángulos señalados con I tenemos $AF = AE$. Luego $AB = AC$ y el triángulo es isósceles.

Uno podría preguntarse por la posición del punto 0, y de hecho se puede probar que está en el círculo circunscrito pero fuera del triángulo. Sin embargo, si se dibuja la figura 42.2, todavía se puede «demostrar» que el triángulo *ABC* es isósceles. El fallo está en que de los dos puntos *E* y *F* uno debe estar dentro y el otro fuera de los respectivos lados del triángulo. Pero esto significa que debemos determinar la posición correcta de *F* con respecto a *A* y *B* y la de *E* con respecto a *A* y *C* antes de comenzar la demostración. Claro

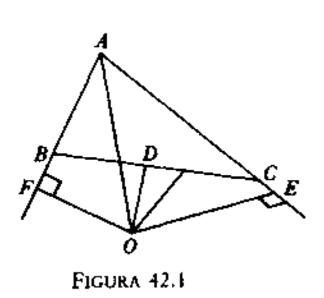


FIGURA 42.1

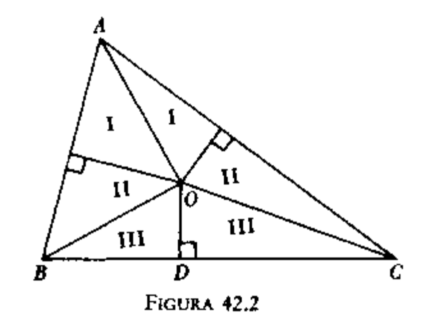


FIGURA 42.2

[3] *Werke*, 8, 193-195 y 200.

está que no habría que confiar en el dibujo para determinar las posiciones precisas de *E* y *F*, pero esto es precisamente lo que Euclides y los matemáticos anteriores a 1800 hacían. Se suponía que la geometría euclídea ofrecía pruebas rigurosas de teoremas sugeridos intuitivamente por las figuras, pero de hecho ofrecía demostraciones intuitivas a partir de figuras dibujadas rigurosamente.

Aunque se alzaron críticas de la estructura lógica de los *Elementos* de Euclides casi desde el momento en que se escribió, no fueron ampliamente conocidas, o se consideraban sus defectos como menores. En general se tomaba a los *Elementos* como un modelo de rigor. No obstante, los trabajos sobre geometrías no euclídeas hicieron a los matemáticos conscientes del alcance de esas deficiencias, exigiendo una actitud especialmente crítica de lo que presuponían al realizar una demostración. El reconocimiento de tales imperfecciones obligó finalmente a los matemáticos a emprender la reconstrucción de los fundamentos de la geometría euclídea y de otras geometrías afectadas de la misma debilidad. Esta actividad se difundió en el último tercio del siglo XIX.

2. Contribuciones a la fundamentación de la geometría proyectiva

En los años setenta del pasado siglo, los trabajos sobre geometría proyectiva en relación con las geometrías métricas revelaron que la fundamental es aquélla (vid. cap. 38). Quizá por esta razón el trabajo de fundamentación comenzó por la geometría proyectiva. Sin embargo, casi todos los autores estaban igualmente interesados en construir las geometrías métricas, ya fuera sobre la base de la geometría proyectiva o independientemente. Por eso los libros y artículos de finales del siglo XIX y comienzos del XX que tratan sobre los fundamentos de la geometría no pueden distribuirse según los diferentes temas.

Los trabajos sobre geometrías no euclídeas habían llevado al convencimiento de que las geometrías son construcciones humanas que se basan en el espacio físico sin ser necesariamente idealizaciones exactas de éste. Lo que implicaba que había que realizar varios cambios importantes en cualquier enfoque axiomático de la geometría. Esto fue reconocido y puesto de relieve por Moritz Pasch (1843-1930), que fue el primero en hacer contribuciones notables a

los fundamentos de la geometría. Sus *Vorlesungen über neuere Geometrie* (1.ª edición en 1882, 2.ª edición, revisada por Max Dehn, en 1926) son una obra pionera en este campo.

Pasch observó que las nociones comunes de Euclides, como las de punto y línea, en realidad no estaban definidas. Decir que un punto es lo que no tiene partes significa bien poco, porque ¿cuál es el significado de «parte»? De hecho, señalaba Pasch, como antes Aristóteles y unos pocos matemáticos más tardíos como Peacock y Boole, algunos conceptos deben quedar sin definición, o si no el proceso de definición sería interminable, o bien la matemática descansaría sobre conceptos físicos. Una vez que se seleccionan ciertos conceptos indefinidos, los demás deben definirse en términos de éstos. Así por ejemplo, en geometría pueden elegirse como términos indefinidos los de punto, recta y plano (Pasch también seleccionaba, en su primera edición, el de congruencia de segmentos de recta). La elección no es única. Como hay términos sin definición, surge la cuestión de qué propiedades de esos conceptos deben usarse para realizar demostraciones con ellos. La respuesta de Pasch es que los axiomas afirman algo acerca de esos términos indefinidos, y que son esos los únicos asertos que pueden utilizarse. Como Gergonne había dicho ya en 1818,[4] los conceptos indefinidos están implícitamente definidos por los axiomas.

En cuanto a éstos, continúa Pasch, aunque algunos pueden ser sugeridos por la experiencia, una vez que se ha seleccionado un conjunto de ellos, debe ser posible realizar todas las demostraciones sin hacer más referencias a la experiencia o al significado físico de los conceptos. Además, los axiomas no son en absoluto verdades auto-evidentes, sino solamente supuestos destinados a proporcionar los teoremas de cualquier geometría particular. En sus *Vorlesungen* (2.ª ed., p. 90), afirma

> ... si la geometría ha de convertirse en una ciencia deductiva genuina, es esencial que la manera en que se realizan inferencias sea independiente tanto del *significado* de los conceptos geométricos como de los diagramas; todo lo que debe considerarse son las relaciones entre los conceptos geométricos aseguradas por las proposiciones y definiciones. Al llevar a cabo una deducción es tan juicioso como útil mantener presente el significado de los conceptos geométricos utilizados, pero *no tiene por qué ser esencial;* de

---

[4] Ann. de Math., 9, 1818-1819, 1.

hecho es precisamente cuando eso se hace necesario cuando se produce un salto en la deducción y (si no es posible colmar la deficiencia modificando el razonamiento) estamos obligados a admitir la inadecuación de las proposiciones invocadas como medio de demostración.

Pasch sí creía que los conceptos y axiomas deben basarse en la experiencia, pero esto era lógicamente irrelevante.

En sus *Vorlesungen* ofrecía axiomas para la geometría proyectiva, pero muchos de esos axiomas o sus análogos fueron igualmente importantes para la axiomatización de las geometrías euclídea y no euclídeas cuando se constituyeron como disciplinas independientes. Como ejemplo, fue el primero en ofrecer un conjunto de axiomas para el orden en que se encuentran los puntos de una recta (que regulan el concepto «estar situado entre»). Tales axiomas deben incluirse también en un conjunto completo para cualquiera de las geometrías métricas. Trataremos más tarde sobre esos conceptos de orden.

Su método para construir la geometría proyectiva consistía en añadir punto, recta y plano del infinito a los puntos, rectas y planos propios. A continuación introducía coordenadas (sobre una base geométrica), utilizando la construcción de Von Staudt y Klein (vid. cap. 35, sec. 3), y finalmente la representación algebraica de las transformaciones proyectivas. Las geometrías euclídea y no euclídeas aparecían como casos especiales sobre una base geométrica distinguiendo los puntos y rectas propios e impropios al estilo de Felix Klein.

Un enfoque más satisfactorio de la geometría proyectiva fue el ofrecido por Peano,[5] Este fue seguido por las obras de Mario Pieri (1860-1904), «*I Principii della geometria di posizione*»,[6] Federigo Enriques (1871-1946), *Lezioni di geometri proiettiva* (1898); Eliakim Hastings Moore (1862-1932);[7] Friedrich H. Schur (1856-1932);[8] Alfred North Whitehead (1861-1947), *The Axioms of Projective Geometry*,[9] y Oswald Veblen (1880-1960) y John W. Young (1879-1932).[10] Estos dos últimos ofrecieron un conjunto completa-

---

[5] *I principii di Geometria*, Turín, Fratelli Bocca, 1889 = Opere Scelte, 2, 56-91.
[6] Memorie della Reale Accademia delle Scienze di Torino, (2), 48, 1899, 1-62.
[7] Amer. Math. Soc. Trans., 3, 1902, 142-158.
[8] Math. Ann., 55, 1902, 265-292.
[9] Cambridge University Press, 1906.
[10] Amer. J. of Math., 1908, 347-378.

mente independiente. Su clásico texto, *Projective Geometry* (2 vols., 1910 y 1918), desarrolla la organización kleiniana de la geometría, partiendo de la geometría proyectiva sobre una base estrictamente axiomática, para particularizar después esa geometría eligiendo diferentes cuádricas absolutas (vid. cap. 38, sec. 3), obteniendo así las geometrías euclídea y no euclídeas. Sus axiomas son lo suficientemente generales como para incluir geometrías con sólo un número finito de puntos, geometrías con sólo puntos racionales, y geometrías con puntos complejos.

Convendría señalar un rasgo más de muchos de los sistemas axiomáticos para la geometría proyectiva y de los que a continuación veremos para la geometría euclídea: algunos de los axiomas de Euclides son axiomas de existencia (vid. cap. 4, sec. 3). Para garantizar la existencia lógica de ciertas figuras, los griegos utilizaban construcciones con regla y compás. Los trabajos de fundamentación del siglo XIX revisaron la noción de existencia, en parte para colmar deficiencias en el manejo que Euclides hacía de este tema, y en parte para ampliar la noción de existencia de manera que la geometría euclídea pudiera incluir puntos, rectas y ángulos no necesariamente construibles con regla y compás. Veremos qué nuevos tipos de axiomas de existencia aparecen en los sistemas que vamos ahora a examinar.

3. Los fundamentos de la geometría euclídea

En su *Sui fondamenti della geometria* (1894), Giuseppe Peano ofrecía un conjunto de axiomas para la geometría euclídea. También él recalcaba que los elementos básicos no deben definirse. Planteaba como principio que debe haber tan pocos conceptos indefinidos como sea posible, que para él eran punto, segmento y movimiento. La inclusión de este último puede parecer un poco sorprendente vista la crítica del uso que hacía Euclides de la superposición; sin embargo, la objeción básica no se refiere al concepto de movimiento sino a la ausencia de una base axiomática adecuada para poder utilizarlo. Un conjunto similar de axiomas fue ofrecido por Pieri,[11] discípulo de Peano, utilizando punto y movimiento como conceptos indefinidos. Otro conjunto, usando línea, segmento y congruencia de seg-

[11] Memorie della Reale Accademia delle Scienze di Torino, (2), 1899, 173-222.

mentos como elementos indefinidos, fue propuesto por Giuseppe Veronese (1854-1917) en sus *Fondamenti di geometria* (1891).

El sistema de axiomas para la geometría euclídea que parece más simple en sus conceptos y proposiciones, el que se mantiene más próximo a Euclides y ha conseguido mayor aceptación, es el debido a Hilbert, que no conocía el trabajo de los italianos. La primera versión apareció en sus *Grundlagen der Geometrie* (1899), pero la revisó varias veces. El sumario que viene a continuación está tomado de la séptima edición (1930) de ese libro. En su utilización de los conceptos indefinidos, cuyas propiedades quedan especificadas únicamente por los axiomas, Hilbert sigue a Pasch: no hay por qué asignar significado explícito alguno a los conceptos no definidos. Esos elementos, punto, recta, plano, etc. podrían reemplazarse, como señalaba Hilbert, por mesas, sillas, jarras de cerveza u otros objetos. Naturalmente, si la geometría trata con «cosas», los axiomas no son en absoluto verdades auto-evidentes, sino que deben ser considerados como arbitrarios aunque de hecho sean sugeridos por la experiencia.

Hilbert enumera primeramente sus conceptos indefinidos: punto, recta, plano, pertenencia de un punto a una recta, pertenencia de un punto a un plano, situación de un punto entre otros dos, congruencia de pares de puntos, y congruencia de ángulos. El sistema de axiomas trata la geometría euclídea plana y sólida conjuntamente y los axiomas aparecen separados en grupos. El primero de ellos contiene los axiomas de existencia:

I. Axiomas de conexión o incidencia

I_1. Para cada dos puntos A y B existe una recta a a la que pertenecen A y B.

I_2. Para cada dos puntos A y B no hay más que una recta a la que pertenezcan A y B.

I_3. En cada recta hay por lo menos dos puntos. Hay por lo menos tres puntos que no pertenecen a la misma recta.

I_4. Para cada tres puntos A, B y C que no están en la misma recta existe un plano α que contiene esos tres puntos. En cada plano hay (por lo menos) un punto.

I_5. Para cada tres puntos A, B y C que no están en la misma recta no hay más que un plano que contenga los tres puntos.

I_6. Si dos puntos de una recta están en un plano α, entonces cada uno de los puntos de la recta está en α.

I_7. Si dos planos α y β tienen un punto A en común, entonces tienen por lo menos otro punto B en común.

I_8. Hay por lo menos cuatro puntos que no están en el mismo plano.

El segundo grupo de axiomas cubre la omisión más seria en el conjunto de Euclides, es decir, la de axiomas acerca del orden en que se encuentran los puntos de una recta:

II. Axiomas de separación

II_1. Si un punto B está entre los puntos A y C, entonces A, B y C son tres puntos diferentes de una recta y B también está entre C y A.

II_2. Para dos puntos cualesquiera A y C hay por lo menos un punto B en la recta AC tal que C está entre A y B.

II_3. Entre tres puntos cualesquiera de una recta, no hay más de uno que esté entre los otros dos.

Los axiomas II_2 y II_3 sirven para hacer la recta infinita.

Definición. Sean A y B dos puntos de una recta a. Se llama segmento AB al par de puntos A, B o B, A. Los puntos entre A y B se llaman puntos del segmento o puntos interiores al segmento. A y B se llaman extremos del segmento. Todos los demás puntos de la recta a se dice que están fuera del segmento.

II_4. (Axioma de Pasch). Sean A, B y C tres puntos no alineados y sea a cualquier recta del plano determinado por A, B y C que no pase por ninguno de esos tres puntos. Si a pasa por algún punto del segmento AB, también debe pasar por algún punto del segmento AC o por algún punto del segmento BC.

III. Axiomas de congruencia

III_1. Si A y B son dos puntos de una recta a y A' es un punto de a o de otra recta a', entonces a un lado (previamente definido) de A' sobre la recta a' se puede encontrar un punto B' tal que el segmento AB es congruente con $A'B'$, lo que se escribe $AB \equiv A'B'$.

III_2. Si $A'B'$ y $A''B''$ son congruentes con AB, entonces $A'B' \equiv A''B''$.

Este axioma limita la afirmación de Euclides «Cosas iguales a una misma cosa son iguales entre sí» a los segmentos de recta.

III_3. Sean AB y BC segmentos sin puntos interiores comunes de una recta a y sean $A'B'$ y $B'C'$ segmentos sin puntos interiores comunes de una recta a'. Si $AB \equiv A'B'$ y $BC \equiv B'C'$, entonces $AC \equiv A'C'$.

Esto equivale a la afirmación de Euclides «Al añadir iguales a iguales se obtienen iguales», aplicada a segmentos de recta.

III_4. Sean $\sphericalangle\ (h,k)$ el ángulo formado por dos semirrectas h y k en un plano α y a' una recta de un plano α' en el que se supone fijado un cierto semiplano de los determinados por a'. Si h' es una semirrecta en a' que parte del punto O', en α' hay una y sólo una semirrecta k' que parte de O' tal que $\sphericalangle\ (h,k)$ es congruente con $\sphericalangle\ (h',k')$ y todos los puntos interiores de $\sphericalangle\ (h',k')$ están en el semiplano fijado de α'. Cada ángulo es congruente consigo mismo.

III_5. Si para dos triángulos ABC y $A'B'C'$ tenemos $AB \equiv A'B'$, $AC \equiv A'C'$ y $\sphericalangle\ BAC \equiv \sphericalangle\ B'A'C'$, entonces $\sphericalangle\ ABC \equiv \sphericalangle\ A'B'C'$.

Este último axioma puede utilizarse para probar que $\sphericalangle\ ACB = \sphericalangle\ A'C'B'$. Se consideran los mismos dos triángulos y las mismas hipótesis. Tomando primero $AC \equiv A'C'$ y después $AB \equiv A'B'$, puede concluirse que $\sphericalangle\ ACB \equiv \sphericalangle\ A'C'B'$ sin más que aplicar el axioma a la nueva reordenación de los términos.

IV. El axioma de las paralelas

Sean a una recta y A un punto que no esté en a. En el plano determinado por a y A hay a lo más una recta que pasa por A sin cortar a a.

La existencia de al menos una recta que pasa por A sin cortar a a puede probarse por otros medios y por tanto no es preciso mencionarlo en el axioma.

V. Axiomas de continuidad

V_1. (Axioma de Arquímedes). Dados dos segmentos cualesquiera AB y CD, en la recta determinada por A y B existen puntos A_1, A_2,

..., A_n tales que los segmentos AA_1, A_1A_2, A_2A_3, ..., $A_{n-1}A_n$ son congruentes con CD y B está entre A y A_n.

V_2. (Axioma de completitud lineal). Los puntos de una recta forman un conjunto que satisface los axiomas I_1, I_2, II, III y V_1 y no puede ampliarse a un conjunto mayor de puntos que siga satisfaciendo esos axiomas.

Este axioma equivale a requerir suficientes puntos en la recta como para que puedan ponerse en correspondencia biunívoca con los números reales. Aunque este hecho había sido utilizado consciente e inconscientemente desde que se empezó a trabajar en geometría con coordenadas, hasta ese momento no se había establecido de su base lógica.

Con estos axiomas Hilbert demostró algunos de los teoremas básicos de la geometría euclídea. Otros autores completaron la tarea de mostrar que toda ella puede deducirse de los axiomas.

El carácter arbitrario de los axiomas de la geometría euclídea, esto es, su independencia con respecto a la realidad física, sacó a la palestra otro problema, el de la consistencia de esa geometría. Mientras se consideró a la geometría euclídea como la verdad acerca del espacio físico, cualquier duda sobre su consistencia parecía infundada. Pero la nueva compresión de los conceptos no definidos y los axiomas exigía que se estableciera aquélla. El problema era aún más vital porque la consistencia de las geometrías no euclídeas se había reducido a la de la geometría euclídea (vid. cap. 38, sec. 4). Poincaré se había ocupado de este asunto en 1898,[12], diciendo que podía creerse en la consistencia de una estructura axiomáticamente fundada, siempre que se pudiera dar de ella una interpretación aritmética. Hilbert mostró que la geometría euclídea es consistente proporcionando tal interpretación.

Para ello identifica (en el caso de la geometría plana) cada punto con un par ordenado (a,b) de números reales,[13] y cada recta con una razón $(u:v:w)$ en la que u y v no son simultáneamente 0. Un punto (a,b) está en una recta $(u:v:w)$ si

$$ua + vb + w = 0$$

La congruencia se interpreta algebraicamente mediante las expre-

---

[12] Monist, 9, 1898, 38.

[13] Estrictamente, usa un conjunto más limitado de números reales.

siones de la geometría analítica para traslaciones y rotaciones; así, dos figuras son congruentes si se puede transformar la una en la otra mediante traslación, simetría con respecto al eje x, y rotación.

Después de haber interpretado aritméticamente cada concepto haciendo ver que esa interpretación satisface los axiomas, Hilbert argumenta que los teoremas deben aplicarse también a la interpretación, puesto que son consecuencia lógica de los axiomas. Si hubiera una contradicción en la geometría euclídea, esta aparecería también en la formulación algebraica de la geometría, que es una extensión de la aritmética. Pero *si la aritmética es consistente,* también lo es entonces la geometría. En aquel momento todavía permanecía abierta la cuestión de la consistencia de la aritmética (vid. cap. 51).

Sería deseable probar que ninguno de los axiomas puede deducirse de algunos o todos los demás, porque en tal caso no habría por qué incluirlo como axioma. Esta noción de independencia fue planteada y discutida por Peano en su artículo de 1894 ya mencionado, e incluso antes en sus *Arithmetices Principia* (1889). Hilbert examinó la independencia de sus axiomas. Sin embargo, en su sistema no es posible mostrar que cada axioma es independiente de todos los demás, ya que el significado de algunos de ellos depende de los precedentes. Lo que sí consiguió probar es que los axiomas de cualquier grupo no pueden deducirse de los axiomas de los otros cuatro grupos. Su método consistió en proporcionar interpretaciones consistentes o modelos que satisfacen los axiomas de cada cuatro grupos sin satisfacer todos los axiomas del quinto grupo.

Estas demostraciones de independencia tenían una relevancia especial en lo que atañe a las geometrías no euclídeas. Para establecer la independencia del axioma de las paralelas, Hilbert ofreció un modelo que satisface los otros cuatro grupos de axiomas pero no ése, utilizando los puntos interiores a una esfera euclídea y ciertas transformaciones que conservan la superficie de la esfera invariante. Así pues, el axioma de las paralelas no puede ser consecuencia de los otros cuatro grupos, porque si lo fuera, el modelo, como parte de la geometría euclídea, poseería propiedades contradictorias en cuanto al paralelismo. La misma demostración prueba que las geometrías no euclídeas son posibles, porque si el axioma euclídeo de las paralelas es independiente de los demás, la negación de ese axioma también debe ser independiente; ya que si fuera una consecuencia, el sistema entero de los axiomas euclídeos contendría una contradicción.

El sistema de axiomas de Hilbert para la geometría euclídea, pu-

blicado por primera vez en 1899, atrajo una considerable atención hacia los fundamentos de la geometría euclídea, y fueron muchos los autores que ofrecieron versiones alternativas utilizando diferentes conjuntos de elementos no definidos o ciertas variaciones en los axiomas. El mismo Hilbert, como ya hemos señalado, hizo varios cambios en su sistema antes de llegar a la versión de 1930. Entre los muchos sistemas alternativos mencionaremos sólo uno: Veblen [14] ofreció un conjunto de axiomas basado en los conceptos no definidos de punto y orden. Mostró que cada uno de sus axiomas es independiente de los demás, y también estableció otra propiedad, conocida con el nombre de categoricidad. Esta noción fue establecida claramente y empleada por primera vez por Edward V. Huntington (1874-1952), en un artículo dedicado al sistema de los números reales [15] (el nombre que él empleaba para esta noción era el de suficiencia). Un conjunto de axiomas P_1, P_2, ..., P_n que conectan un conjunto de símbolos no definidos S_1, S_2, ..., S_m se dice que es categórico si entre los elementos de dos colecciones cualesquiera, cada una de las cuales contiene símbolos no definidos y satisface los axiomas, se puede establecer una correspondencia biunívoca para los conceptos no definidos que preserve las relaciones establecidas por los axiomas; esto es, ambos sistemas son isomorfos. La categoricidad significa así que las diferentes interpretaciones del sistema de axiomas difieren únicamente en lenguaje. Esta propiedad no se cumpliría, por ejemplo, si se omitiera el axioma de las paralelas, ya que entonces la geometría euclídea y la hiperbólica serían interpretaciones no isomorfas del conjunto reducido de axiomas.

La categoricidad implica otra propiedad que Veblen llamaba disyuntiva y que se conoce ahora como completitud. Se dice que un sistema de axiomas es completo si es imposible añadir otro axioma que sea independiente del conjunto dado y consistente con él (sin introducir nuevos conceptos primitivos). La categoricidad implica la completitud, ya que si un conjunto de axiomas A fuese categórico y no completo se podría introducir un axioma S tal que tanto S como no-S fueran consistentes con el conjunto A. Y como el conjunto original A es categórico, habría interpretaciones isomorfas de A con S y de A con no-S. Pero eso es imposible, porque deben

[14] Amer. Math. Soc. Transl., 5, 1904, 343-384.
[15] Amer. Math. Soc. Transl., 3, 1902, 264-279.

cumplirse las proposiciones correspondientes de ambas interpretaciones, y *S* se aplica a una interpretación y no-*S* a la otra.

4. Otros trabajos de fundamentación

La nitidez alcanzada en los axiomas para la geometría euclídea sugirió investigaciones paralelas para las varias geometrías no euclídeas. Una de las características notables de los axiomas de Hilbert es que los axiomas para la geometría hiperbólica pueden obtenerse inmediatamente reemplazando el axioma euclídeo de las paralelas por el axioma de Lobatchevski-Bolyai, permaneciendo inalterados los demás.

Para obtener axiomas para la geometría simple o doblemente elíptica, no basta abandonar el axioma euclídeo de las paralelas para sustituirlo por un axioma que haga que dos rectas cualesquiera tengan un punto en común (elíptica simple) o al menos un punto en común (elíptica doble), y hay que cambiar también otros axiomas. La recta de esas geometrías no es infinita, sino que tiene las propiedades de una circunferencia. Así pues, hay que reemplazar los axiomas de orden de la geometría euclídea por otros que describan las relaciones de orden entre los puntos de una circunferencia. Se han propuesto varios sistemas de axiomas de este tipo. George B. Halsted (1853-1922), en su *Rational Geometry*,[16] y John R. Kline (1891-1955)[17] ofrecieron bases axiomáticas para la geometría doblemente elíptica, y Gerhard Hessenberg (1874-1925)[18] propuso un sistema de axiomas para la geometría elíptica simple.

Otro tipo de investigaciones sobre la fundamentación de la geometría son las referidas a las consecuencias de la negación u omisión de uno o más axiomas de un determinado conjunto. El mismo Hilbert ya había hecho esto en sus demostraciones de independencia, ya que la esencia de tales demostraciones consiste en la construcción de un modelo o interpretación que satisfaga todos los axiomas excepto aquél cuya independencia se pretende establecer. El ejemplo más significativo de axioma negado es, naturalmente, el de las paralelas. Pero también se han obtenido resultados interesantes eliminan-

---

[16] 1904, 212-247.
[17] Annals of Math., (2), 18, 1916-1917, 31-44.
[18] Math. Ann., 61, 1905, 173-184.

do el axioma de Arquímedes, que en el sistema de Hilbert aparece como V_1. La geometría resultante se llama no-arquimediana; en ella hay pares de segmentos tales que al multiplicar uno de ellos por cualquier número entero, por grande que sea, no se consigue un segmento mayor que el otro. En sus *Fondamenti di geometria*, Giuseppe Veronese construyó una geometría no arquimediana, mostrando también que los teoremas de esa geometría se aproximan tanto como se quiera a los de la geometría euclídea.

Max Dehn (1878-1952) también obtuvo [19] muchos teoremas interesantes omitiendo el axioma de Arquímedes. Por ejemplo, hay una geometría en la que la suma de los ángulos de un triángulo es dos rectos, existen triángulos semejantes no congruentes y se pueden trazar infinitas paralelas a una recta dada pasando por un punto también dado.

Hilbert había señalado que para la construcción de la teoría de las áreas en el plano no se necesitaba el axioma de continuidad V_2. Para el espacio, sin embargo, Max Oehn probó [20] la existencia de poliedros que tienen el mismo volumen y no se pueden descomponer en partes mutuamente congruentes (incluso añadiéndoles poliedros congruentes). Así pues, en tres dimensiones el axioma de continuidad es necesario.

La fundamentación de la geometría euclídea fue enfocada de manera completamente diferente por algunos matemáticos. La geometría, como sabemos, había caído en desgracia cuando los matemáticos se dieron cuenta de que habían aceptado ciertos hechos inconscientemente, basándose en la intuición, y sus supuestas demostraciones eran consiguientemente incompletas. El peligro de que esto pudiera repetirse una y otra vez les hizo creer que la única base sólida para la geometría sería la aritmética. La forma en que podían erigir esa base estaba clara. De hecho, Hilbert había ofrecido una interpretación aritmética de la geometría euclídea. Lo que había que hacer ahora, para la geometría plana por ejemplo, era no interpretar un punto como un par de números (x,y), sino *definir* como punto el par de números, como recta la razón entre tres números $(u:v:w)$, definir la pertenencia del punto (x,y) a la recta $(u:v:w)$ por la satisfacción de la ecuación $ux + vy + w = 0$, definir como circunferencia el conjunto de los (x,y) que satisfacen una ecuación del tipo $(x - a)^2$

[19] Math. Ann., 53, 1900, 404-439.
[20] Math. Ann., 55, 1902, 465-478.

$+ (y - b)^2 = r^2$, y así sucesivamente. Con otras palabras, habría que utilizar los equivalentes analíticos de las nociones puramente geométricas como definiciones de los conceptos geométricos, y métodos algebraicos para probar los teoremas. Como la geometría analítica contiene en forma algebraica la contrapartida exacta de cuanto existe en geometría euclídea, no había problema en cuanto a la obtención de la fundamentación aritmética. De hecho, el trabajo técnico necesario estaba ya hecho, incluso para la geometría euclídea n-dimensional, por ejemplo en la *Ausdehnungslehre* de Grassmann; y el mismo Grassmann había propuesto que su obra sirviera como fundamentación para la geometría euclídea.

5. Algunas cuestiones abiertas

Las investigaciones críticas sobre la geometría se extendieron más allá de la reconstrucción de los fundamentos. Hasta entonces se habían utilizado libremente las curvas. Las más simples, como la elipse, tenían definiciones geométricas y analíticas seguras. Pero muchas curvas se habían introducido únicamente por medio de ecuaciones y funciones. La rigorización del análisis había supuesto no sólo la ampliación del concepto de función, sino también la construcción de funciones muy peculiares, como las funciones continuas sin derivada en ningún punto. Que estas funciones atípicas fueran turbadoras desde el punto de vista geométrico es fácil de entender. La curva que representa el ejemplo de Weierstrass de una función que es continua en todo punto pero no es diferenciable en ninguno no se ajusta evidentemente al concepto usual, porque la falta de derivada significa que esa curva no puede tener tangente en ningún punto. La cuestión que surgió entonces es si las representaciones geométricas de tales funciones son curvas y, más en general, la de qué debe entenderse por curva.

Jordan dio como definición de curva [21] el conjunto de puntos representados por las funciones continuas $x = f(t)$, $y = g(t)$, para $t_0 \leq t \leq t_1$. Para algunos propósitos Jordan quería restringir sus curvas de manera que no poseyeran puntos múltiples, y requirió entonces que $f(t) \neq f(t')$ o $g(t) \neq g(t')$ para t y t' entre t_0 y t_1, es

[21] *Cours d'Analyse*, 1.ª ed., vol. 3, 1887, 593; 2.ª ed., vol. 1, 1893, 90.

decir, que para cada (x,y) de la curva hubiese un solo t. Tales curvas reciben el nombre de curvas de Jordan.

En este mismo tratado presentó la noción de curva cerrada,[22] requiriendo únicamente que $f(t_0) = f(t_1)$ y $g(t_0) = g(t_1)$, y estableció el teorema según el que una curva cerrada divide al plano en dos partes, una interior y otra exterior. Dos puntos de la misma región pueden unirse mediante un camino poligonal que no corta a la curva, y dos puntos situados en diferentes regiones no pueden unirse mediante ninguna línea poligonal o curva continua que no corte a la curva cerrada simple. El teorema es más potente de lo que parece a primera vista, ya que una curva cerrada simple puede tener un aspecto muy intrincado. De hecho, como las funciones $f(t)$ y $g(t)$ sólo tienen que ser continuas, resulta implicada toda la variedad de posibles complicaciones que pueden aparecer en las funciones continuas. El mismo Jordan y muchos matemáticos distinguidos propusieron demostraciones incorrectas del teorema. La primera demostración rigurosa se debe a Veblen [23].

No obstante, la definición de Jordan de curva, aun siendo satisfactoria para muchas aplicaciones, era demasiado amplia. En 1890 Peano descubrió [24] que una curva satisfaciendo la definición de Jordan puede pasar por todos los puntos de un cuadrado al menos una vez. Ofreció una descripción aritmética detallada de la correspondencia entre los puntos del intervalo [0,1] y los puntos del cuadrado, especificando dos funciones $x = f(t)$, e $y = g(t)$, unívocas y continuas para $0 \leqslant t \leqslant 1$ y tales que x e y toman los valores correspondientes a cada punto del cuadrado unidad. Sin embargo, la correspondencia de (x,y) a t no es unívoca, ni es continua. Una correspondencia biunívoca continua de los valores de t con los valores de (x,y) es imposible; o sea, que $f(t)$ y $g(t)$ no pueden ser entonces simultáneamente continuas. Quien probó esto fue Eugen E. Netto (1846-1919) [25].

La interpretación geométrica de la curva de Peano fue ofrecida por Arthur M. Schoenflies (1853-1928) [26] y E. H. Moore [27]. Primeramente se transforma el segmento [0,1] en los nueve segmentos

[22] 1.ª ed., p. 593; 2.ª ed., p. 98.

[23] Amer. Math. Soc. Trans., 6, 1905, 83-98, y 14, 1913, 65-72.

[24] Math. Ann. 36, 1890, 157-160 = *Opere Scelte*, 1, 110-115.

[25] Jour. für Math., 86, 1879, 263-268.

[26] Jahres. der Deut. Math.-Verein, 82, 1900, 121-125.

[27] Amer. Math. Soc. Trans., 1, 1900, 72-90.

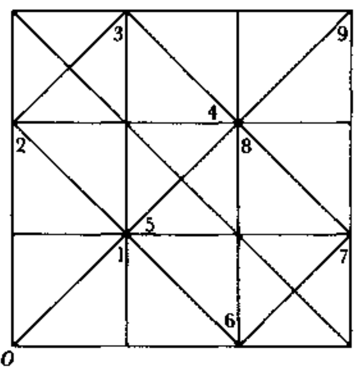


FIGURA 42.3

presentados en la figura 42.3, y a continuación en cada subcuadrado se descompone la diagonal en él contenida siguiendo el mismo método, pero haciendo que la transición desde cada subcuadrado al siguiente sea continua, y se repite el proceso *ad infinitum*, consiguiendo en el límite cubrir todo el cuadrado original. Ernesto Cesàro (1859-1906) [28] presentó la forma analítica de las f y g de Peano.

Hilbert ofreció otro ejemplo [29] de aplicación continua del segmento unidad sobre el cuadrado. Divídanse el segmento unidad y el cuadrado en cuatro partes iguales, tal como aparecen en la figura 42.4;

Atravesemos ahora cada subcuadrado de manera que el camino mostrado en la figura corresponda al segmento unidad. Dividamos a continuación el cuadrado unidad en 16 subcuadrados, numerándolos en la forma que muestra la figura 42.5, y unamos los centros de esos 16 subcuadrados en la forma también indicada.

El proceso continúa dividiendo cada subcuadrado en cuatro partes, y numerando éstas de manera que podamos recorrer todo el conjunto siguiendo un camino continuo. La curva deseada es el límite de las sucesivas poligonales obtenidas en cada etapa. Como los subcuadrados y las partes del segmento unidad se contraen simultáneamente hacia puntos al proseguir la subdivisión, puede verse intuitivamente que cada punto del segmento unidad se transforma así en un solo punto del cuadrado. De hecho, si fijamos un punto del

---

[28] Bul. des Sci. Math., (2), 21, 1897, 257-266.

[29] Math. Ann., 38, 1891, 459-460 = *Ges. Abh.*, 3, 1-2.

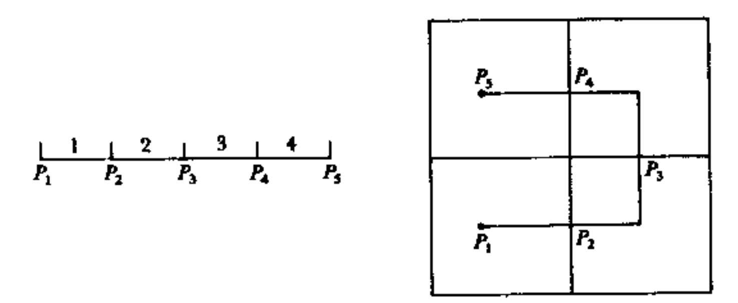


FIGURA 42.4

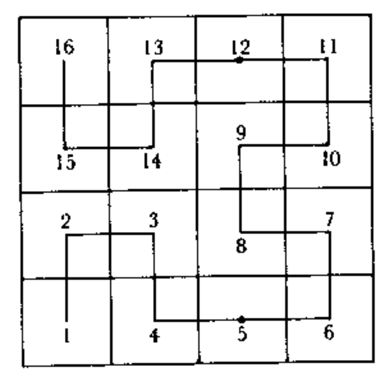


FIGURA 42.5

segmento unidad, digamos $t = 2/3$, entonces la imagen de ese punto es el límite de las sucesivas imágenes de $t = 2/3$ que aparecen en las diferentes poligonales.

Estos ejemplos muestran que la definición de curva sugerida por Jordan no es satisfactoria, ya que una curva, según esa definición, puede llenar un cuadrado. La cuestión de qué debe entenderse por curva permanecía por tanto abierta. Felix Klein señaló en 1898 que no había nada más oscuro que la noción de curva [30]. Y fueron los topólogos quienes afrontaron el problema (vid. cap. 50, sec. 2).

Más allá del problema de qué debe entenderse por curva, la ex-

[30] Math. Ann., 50, 1898, 586.

tensión del análisis a las funciones sin derivada planteó también la cuestión de la longitud que se supone debe tener una curva. La acostumbrada fórmula del cálculo es

$$L = \int_a^b (1 + y'^2)^{1/2}\, dx,$$

donde $y = f(x)$ exige, desde su mismo planteamiento, la existencia de derivada, y no puede aplicarse a las funciones no diferenciables. Du Bois Reymond, Peano, Ludwig Scheefer (1859-1885) y el mismo Jordan realizaron varios esfuerzos por generalizar el concepto de longitud de una curva, utilizando definiciones de integral más generales o conceptos geométricos. La definición más general se alcanzó en relación con la noción de medida, que examinaremos en el capítulo 44.

Una dificultad similar fue señalada para el concepto de área de una superficie. La idea más extendida en los textos del siglo XIX consistía en inscribir en la superficie un poliedro con caras triangulares, tomando como área de la superficie el límite de las sumas de las áreas de esos triángulos cuando sus lados tienden a O. Analíticamente, si la superficie viene representada por

$$x = \phi(u, v), \qquad y = \psi(u, v), \qquad z = \chi(u, v)$$

la fórmula para el área de la superficie se convierte en

$$\int_D \sqrt{A^2 + B^2 + C^2}\, du\, dv,$$

donde A, B y C son los jacobianos de y y z, x y z, y, x e y, respectivamente. De nuevo se planteaba la cuestión de qué hacer cuando x, y o z no poseen derivadas. Para complicar aún más la situación, H. A. Schwarz propuso en una carta a Hermite [31] un ejemplo en el que una elección determinada de los triángulos del poliedro inscrito conducía a un área infinita incluso para un cilindro ordinario [32]. La teoría del área de una superficie tuvo que ser también reconsiderada en relación con la noción de medida.

---

[31] *Ges. Math. Abh.*, 2, 309-311.

[32] Puede encontrarse este ejemplo en J. Pierpont, *The theory of functions of a real variable*, Dover (reimpresión), 1959, vol. 2, p. 26.

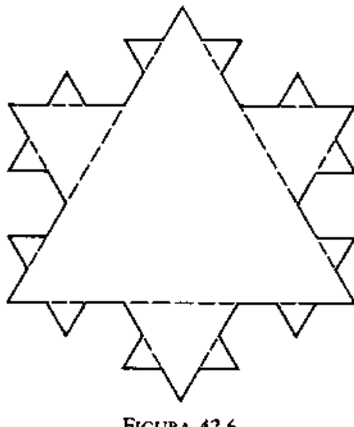


FIGURA 42.6

Hacia 1900 nadie había probado que una curva cerrada plana, tal como la habían definido Jordan y Peano, encerrara un área. Helge von Koch (1870-1924) complicó aún más el problema del área ofreciendo un ejemplo de una curva continua aunque no diferenciable, con perímetro infinito, que limitaba un área finita.[33] Se parte de un triángulo equilátero *ABC* (fig. 42.6) de lado 3*s*. Sobre el tercio central de cada lado se construye un triángulo equilátero de lado *s* y se elimina la base de esos nuevos triángulos, que serán tres. Sobre cada lado de longitud *s* de la nueva figura se construye, hacia el exterior, tomando como base el tercio central, un triángulo equilátero de lado $s/3$, y se eliminan las bases de todos esos nuevos triángulos, que serán 12. Sobre el tercio central de cada lado de la figura resultante se construye ahora un triángulo equilátero de lado $s/9$, de los que habrá 48. Los perímetros de los sucesivos polígonos son 9*s*, 12*s*, 16*s*, ... y tienden por tanto a infinito. Sin embargo, el área de la figura limitada por esos polígonos se mantiene acotada. En efecto, por la bien conocida fórmula del área de un triángulo equilátero en función de su lado, es decir $(b^2/4)\sqrt{3}$ si la longitud del lado es *b*, el área del triángulo original es $[(3s)^2/4]\sqrt{3}$. El área de los tres primeros triángulos añadidos es $3 \cdot (s^2/4)\sqrt{3}$. Como la longitud de los lados de los triángulos añadidos en la siguiente etapa es $s/3$ y hay 12, el área

[33] Acta Math., 30, 1906, 145-176.

total añadida es $12 \cdot (s/3)^2 \sqrt{3}/4 = (s^2/3)\sqrt{3}$. La suma de las áreas es

$$S = \frac{9s^2}{4}\sqrt{3} + \frac{3}{4}s^2\sqrt{3} + \frac{s^2}{3}\sqrt{3} + \frac{4s^2}{27}\sqrt{3} + \ldots$$

Se trata de una serie geométrica (excluyendo el primer sumando) de razón 4/9, cuya suma es

$$S = \frac{9s^2}{4}\sqrt{3} + \frac{(3/4)s^2\sqrt{3}}{1 - 4/9} = \frac{18}{5}s^2\sqrt{3}.$$

Las curvas de Peano y Hilbert también plantearon la cuestión de qué debe entenderse por dimensión. El cuadrado es por sí mismo bidimensional, pero como imagen continua de una curva, tendría que ser unidimensional. Además, Cantor había mostrado que los puntos de un segmento de recta pueden ponerse en correspondencia biunívoca con los puntos de un cuadrado (cap. 41, sec. 7). Aunque esa correspondencia no es continua en ninguno de los dos sentidos, muestra que la dimensión no es una cuestión de cantidad de puntos. Tampoco es cuestión del número de coordenadas necesarias para fijar la posición de un punto, como habían pensado Riemann y Helmholtz, ya que la curva de Peano asigna un único (x,y) a cada valor de t.

A la luz de estas dificultades puede verse que la rigorización de la geometría no bastaba para responder a todos los problemas que se habían planteado. Muchos de ellos fueron resueltos por los topólogos y analistas del presente siglo. El mismo hecho de que continuaran surgiendo problemas acerca de los conceptos fundamentales muestra una vez más que las matemáticas no crecen como una estructura lógica. Los avances en nuevos campos, e incluso el perfeccionamiento de otros más antiguos, revelan nuevos e insospechados defectos. Más allá de la resolución de los problemas que conciernen a curvas y superficies tenemos todavía que ver si se alcanzó la última etapa en cuanto al rigor con los trabajos de fundamentación en análisis, el sistema de los números reales y la geometría básica.

Bibliografía

Becker, Oskar: *Grundlagen der Mathematik in geschichtlicher Entwicklung,* Karl Alber, 1954, 199-212.

Enriques, Federigo: «Prinzipien der Geometrie». *Encyk. der Math. Wiss.,* B. G. Teubner, 1907-1910, III AB1, 1-129.

Hilbert, David: *Grundlagen der Geometrie, 7.ª* ed., B. G. Teubner, 1930. Hay versión castellana, D. Hilbert, *Fundamentos de la Geometría,* Madrid, C.S.I.C., 1991.

Pasch, M. y M. Dehn: *Vorlesungen über neuere Geometrie, 2.ª* ed., Julius Springer, 1926, 185-271.

Peano, Giuseppe: *Opere scelte,* 3 vols. Edizioni Cremonese, 1957-1959.

Reichardt, Hans: *C. F. Gauss, Leben und Werke,* Haude und Spenersche, 1960, 111-150.

Schmidt, Arnold: «Zu Hilberts Grundlegung der Geometrie», en las *Gesammelte Abhandlungen* de Hilbert, 2, 404-414.

Singh, A. N.: «The Theory and Construction of Non-Differentiable Functions», en E. W. Hobson: *Squaring the Circle and Other Monographs,* Chelsea (reimpresión), 1953.

Capítulo 43
LA MATEMATICA EN TORNO A 1900

> No he vacilado en llamar al siglo XIX, en el Congreso de Matemáticos de París en 1900, el siglo de la teoría de funciones.
>
> VITO VOLTERRA

1. Las principales características de los desarrollos del siglo XIX

Tanto en el XIX como en los dos siglos precedentes los avances en matemáticas trajeron consigo profundos cambios, apenas perceptibles en el desarrollo de un año a otro, pero trascendentes en sí mismos y en su efecto sobre futuras investigaciones. La vasta expansión de los temas y la incorporción de nuevos campos, tanto como la extensión de otros más antiguos, son evidentes. El álgebra recibió un impulso totalmente nuevo con Galois; la geometría recobró vitalidad y se vio profundamente alterada por la introducción de las geometrías no euclídeas y el resurgimiento de la geometría proyectiva; la teoría de números se transformó en teoría analítica de números; y el análisis se amplió extraordinariamente con la introducción de la teoría de funciones complejas y la expansión de las ecuaciones diferenciales ordinarias y en derivadas parciales. Desde el punto de vista del desarrollo técnico, la teoría de funciones de variable compleja fue la más significativa de las nuevas creaciones. Pero desde el punto de vista de la importancia intelectual y de sus efectos más profundos sobre la naturaleza de la matemática, el desarrollo más renovador fue la geometría no euclídea. Como veremos, sus efectos

fueron mucho más revolucionarios de lo que hasta ahora hemos indicado. El círculo en el que los estudios matemáticos parecían estar encerrados a comienzos del siglo se rompió por muchos puntos, y las matemáticas proliferaron en un centenar de ramas. El torrente de nuevos resultados contradijo agudamente la opinión dominante a finales del siglo XVIII de que el caudal de las matemáticas había quedado exhausto.

La actividad matemática también se expandió en otros aspectos durante el siglo XIX. El número de matemáticos se incrementó enormemente como consecuencia de la democratización de la enseñanza. Aunque Alemania, Francia e Inglaterra seguían siendo los centros más importantes, Italia reapareció en la arena, y los Estados Unidos, con Benjamín Peirce, G. W. Hill y Josiah Williard Gibbs entraron en ella por primera vez. En 1863 se fundó en los USA la Academia Nacional de Ciencias. Sin embargo, a diferencia de la Royal Society de Londres, o las Academias de Ciencias de París o Berlín, la Academia Nacional americana no ha sido un lugar de encuentros científicos en los que se presentaran y revisaran los nuevos resultados, aunque sí publica una revista, los *Proceedings of the Academy*. Se organizaron otras sociedades matemáticas (vid. cap. 26, sec. 6), en las que los investigadores podían reunirse y presentar sus artículos, y que patrocinaban diferentes revistas. A fines de siglo el número de éstas dedicadas parcial o enteramente a la investigación matemática había aumentado hasta unas 950. En 1897 comenzó la costumbre de mantener un congreso internacional cada cuatro años.

Junto a la explosión de la actividad matemática se produjo una novedad menos agradable. Las diferentes disciplinas se hicieron autónomas, dando lugar cada una de ellas a su propia terminología y metodología. La búsqueda de cualquier resultado impuso el estudio de problemas más especializados y más difíciles, que requerían más ideas, y más ingeniosas, una fértil inspiración, y demostraciones más intrincadas. Para realizar cualquier avance, los matemáticos se vieron obligados a adquirir una sólida fundamentación teórica y gran habilidad técnica. La especialización se hizo notable en los trabajos de Abel, Jacobi, Galois, Poncelet y muchos otros. Aunque se puso algún acento en la interrelación entre diferentes ramas con nociones como las de grupo, transformación lineal o invariancia, el efecto general fue el de una separación en numerosos campos distintos y alejados entre sí. A Felix Klein le pareció en 1893 que se podría

superar la especialización y divergencia entre las diferentes ramas mediante los conceptos que acabamos de mencionar, pero su esperanza fue vana. Cauchy y Gauss fueron los últimos en conocer la matemática como un todo, aunque Poincaré y Hilbert fueron también matemáticos casi universales.

Desde el siglo XIX los matemáticos trabajan normalmente sólo en algún área de la matemática, subrayando naturalmente cada uno la importancia de la propia sobre las restantes. Sus publicaciones ya no son para un público amplio, sino para un círculo restringido de colegas. La mayoría de los artículos ya no contienen ninguna indicación de su conexión con otros problemas de la matemática, son difícilmente accesibles para muchos matemáticos, e incomprensibles para los no iniciados.

Más allá de sus logros en temas concretos, el siglo XIX se caracteriza por la reintroducción del rigor en las demostraciones. Dejando de lado lo que los matemáticos puedan haber pensado individualmente acerca de la solidez de sus resultados, el hecho es que desde el año 200 a. de C. hasta 1870 aproximadamente, casi toda la matemática se sostuvo sobre una base empírica y pragmática, perdiéndose de vista el concepto de demostración deductiva a partir de axiomas explícitos. Una de las revelaciones más sorprendentes de la historia de las matemáticas es precisamente esa relajación en el rigor deductivo durante los dos milenios en los que su contenido se extendió tan ampliamente. Aunque cabe reconocer algunos esfuerzos más tempranos por rigorizar el análisis, en particular los de Lagrange (cap. 19, sec. 7), la actitud dominante era la representada por Lacroix (cap. 26, sec. 3). Los trabajos de Fourier pueden poner los pelos de punta a un analista de nuestros días; en cuanto a Poisson, la derivada y la integral no eran para él sino abreviaturas para un cociente de diferencias o una suma finita. El movimiento iniciado por Bolzano y Cauchy para asentar los fundamentos surgió indudablemente de la preocupación por la cantidad rápidamente creciente de matemáticas que descansaban sobre los frágiles fundamentos del cálculo, y se vio acelerado por el descubrimiento de Hamilton de los cuaterniones no conmutativos, que desafiaba los principios acríticamente aceptados acerca de los números. Pero más turbadora aún fue la creación de las geometrías no euclídeas, que no sólo destruyó la mismísima noción de la auto-evidencia de los axiomas y su aceptación demasiado superficial, sino que reveló insuficiencias en demostraciones que habían sido consideradas como las más sólidas de toda la matemática.

Los matemáticos se dieron cuenta de que habían pecado de credulidad confiando en la intuición.

Hacia 1900 parecía alcanzado el objetivo de proporcionar una fundamentación rigurosa a la matemática, y los matemáticos casi se envanecían de semejante logro. En su intervención ante el Segundo Congreso Internacional en París,[1] Poincaré exultaba: «¿Hemos alcanzado por fin el rigor absoluto? En cada etapa de su evolución nuestros predecesores también creyeron haberlo alcanzado. Y si ellos se equivocaron, ¿no estaremos equivocados nosotros también?... En el análisis actual, no obstante, si tenemos el cuidado de ser rigurosos, sólo hay silogismos o invocaciones a la intuición de número puro, que no pueden engañarnos. Ahora sí se puede decir que se ha alcanzado el rigor absoluto.» Cuando se consideran los resultados clave en la fundamentación del sistema numérico y la geometría, y la erección del análisis sobre la base de ese sistema numérico, se entienden las razones para el envanecimiento. Las matemáticas contaban ahora con unos fundamentos que prácticamente cualquiera estaría dispuesto a aceptar.

La formulación precisa de conceptos básicos como los de número irracional, continuidad, integral, y derivada no fue sin embargo recibida entusiásticamente por todos los matemáticos. Muchos de ellos no entendían el nuevo lenguaje $\varepsilon - \delta$, y consideraban esas definiciones precisas como caprichosas e innecesarias para la comprensión de las matemáticas, e incluso para el rigor en las demostraciones, sintiendo que bastaba con la intuición, a pesar de las sorpresas de las funciones continuas sin derivada, las curvas que llenan un espacio, o las curvas sin longitud. Emile Picard, por ejemplo, decía a propósito del rigor en las ecuaciones en derivadas parciales: «... el verdadero rigor es productivo, y se distingue en eso de aquel otro rigor que es puramente forma y enojoso, y que arroja una sombra sobre cada uno de los problemas que toca.»[2]

A pesar de que la geometría también se ha visto sometida a la rigorización, una de las consecuencias de este movimiento fue que números y análisis se elevaran por encima de ella. El reconocimiento por los matemáticos, durante y después de la creación de las geo-

---

[1] *Comp. Rendus du Deuxième Congres Internat. des Math.*, 1900, pub. 1902, 121-122.

[2] Amer. Math. Soc. Bull., 11, 1904-1905, 417; ver también cap. 40, sec. 7 y cap. 41, sec. 5 y 9.

metrías no euclídeas, de que se habían dejado llevar inconscientemente por la intuición al aceptar las demostraciones de la geometría euclídea, les hizo temer que pudiera seguir siendo así en cualquier razonamiento geométrico, y preferir una matemática construida sobre los números. Muchos estuvieron a favor de ir más allá y construir toda la geometría sobre los números, lo que podía hacerse por medio de la geometría analítica en la forma descrita anteriormente. Así, la mayoría de los matemáticos hablaba de la aritmetización de las matemáticas, aunque habría sido más preciso hablar de la aritmetización del análisis. Donde Platón pudo decir «Dios geometriza eternamente», Jacobi, ya a mediados de siglo, dijo: «Dios aritmetiza permanentemente». En el Segundo Congreso Internacional, Poincaré aseguraba: «Actualmente sólo quedan en análisis los enteros y sistemas finitos e infinitos de enteros, interrelacionados por una red de relaciones de igualdad o desigualdad. La matemática, como decimos, ha sido aritmetizada.» Si Pascal había dicho: *«Tout ce qui passe la Géométrie nous passe»*,[3] los matemáticos de 1900 preferían decir: *«Tout ce qui passe l'Arithmétique nous passe.»*

La construcción de fundamentos lógicos para las matemáticas, dejando a un lado la preferencia por una base aritmética o geométrica, completó otra etapa en la ruptura con la metafísica. La vaguedad en los fundamentos y justificaciones de las inferencias matemáticas se había ocultado en el siglo XVIII y a comienzos del XIX con alusiones a argumentos metafísicos que, aunque nunca se explicitaban, eran mencionados como respaldo para aceptar las matemáticas. La axiomatización de los números reales y la geometría proporcionó a la matemática una base clara, independiente y autosuficiente. Ya no era preciso recurrir a la metafísica. Como dijo Lord Kelvin, «la matemática es la única metafísica aceptable».

La rigorización de la matemática pudo cubrir una necesidad del siglo XIX, pero también nos enseña algo acerca de su desarrollo. La estructura lógica recién implantada se suponía que garantizaba la solidez de las matemáticas, pero esa garantía tenía algo de impostura: ni un solo teorema de aritmética, álgebra o geometría euclídea fue refutado por el nuevo rigor, y los teoremas de análisis sólo tuvieron que ser formulados más cuidadosamente. De hecho, todo lo que hicieron las nuevas estructuras axiomáticas fue confirmar lo que los matemáticos sabían ya; los axiomas permitían inferir los teoremas

[3] «Todo lo que trasciende la geometría trasciende nuestra comprensión».

existentes, pero no los determinaban. Lo que significa que la matemática descansa, no en la lógica, sino en intuiciones profundas. El rigor, como señalaba Jacques Hadamard, sólo sanciona las conquistas de la intuición; o como decía Hermann Weyl, la lógica es la higiene que practica el matemático para mantener sus ideas fuertes y saludables.

2. El movimiento axiomático

La rigorización de la matemática se consiguió axiomatizando sus diferentes ramas. La esencia de un desarrollo axiomático, de acuerdo con el modelo que hemos descrito en los capítulos 41 y 42, consiste en partir de unos términos no definidos cuyas propiedades quedan especificadas por los axiomas; el objetivo de todo el trabajo es obtener los teoremas como consecuencia de los axiomas. Además, hay que establecer para cada sistema la independencia, consistencia y categoricidad de los axiomas, nociones que ya hemos examinado en los dos capítulos precedentes.

A comienzos del siglo XX el método axiomático no sólo permitió establecer los fundamentos lógicos de muchas ramas, viejas y nuevas, de la matemática, sino que reveló precisamente qué suposiciones subyacen bajo cada una de ellas e hizo posible la comparación y clarificación de las relaciones entre esas diferentes ramas. Hilbert estaba entusiasmado con la validez de este método. Analizando el estado presumiblemente perfecto que la matemática había alcanzado fundamentando cada una de sus ramas sobre bases axiomáticas sólidas, señalaba:[4]

> De hecho, el método axiomático ha sido y sigue siendo la ayuda más conveniente e indispensable para cualquier investigación exacta en no importa qué dominio; es lógicamente inobjetable y al mismo tiempo fructífero, y garantiza una completa libertad de investigación. Proceder axiomáticamente significa pues pensar con conocimiento de lo que uno se trae entre manos. Mientras que antes, sin el método axiomático, se actuaba ingenuamente creyendo en ciertas relaciones como en un dogma, el enfoque axiomático aparta esa ingenuidad sin privarnos por eso de las ventajas de la creencia.

---

[4] Abh. Math. Seminar Hamburger Univ., 1, 1922, 157-177 = *Ges. Abh.*, 3, 157-177.

También en el último apartado de su «Axiomatisches Denken»,[5] defendía este método:

> Todo lo que puede ser objeto del pensamiento matemático, en cuanto esté madura la construcción de una teoría, cae bajo el método axiomático y por tanto directamente dentro de la matemática. Ahondando en capas cada vez más profundas de los axiomas... podemos obtener percepciones más profundas del pensamiento científico y constatar la unidad de nuestro conocimiento. Gracias especialmente al método axiomático, la matemática parece llamada a desempeñar un papel conductor en todo conocimiento.

La posibilidad de explorar nuevos problemas omitiendo, negando o variando de alguna otra manera los axiomas de sistemas establecidos atrajo a muchos matemáticos. Esta actividad, y la construcción de bases axiomáticas para las diferentes ramas de la matemática se conocen como movimiento axiomático. Sigue siendo todavía una actividad muy practicada. Parte de su atractivo se explica por el hecho de que, una vez establecidas bases axiomáticas sólidas para las ramas más importantes, variaciones del tipo que acabamos de describir son relativamente fáciles de introducir y explorar. Sin embargo, cualquier nuevo desarrollo en la matemática ha atraído siempre a un cierto número de personas que buscan campos abiertos a la exploración o están sinceramente convencidos de que el futuro de la matemática reside en ese área particular.

3. La matemática como creación del hombre

Desde el punto de vista del futuro desarrollo de las matemáticas, el resultado más significativo del siglo fue la obtención de un enfoque apropiado de las relaciones de las matemáticas con la naturaleza. Aunque no hemos entrado en las opiniones sobre las matemáticas de muchos de los autores cuyos trabajos hemos descrito, sí hemos dicho que los griegos, Descartes, Newton, Euler y muchos otros creían que las matemáticas proporcionan la descripción precisa de los fenómenos reales y que consideraban sus trabajos como desvelamientos de la estructura matemática del universo. Los matemáticos trataban con abstracciones, pero éstas no eran más que las formas

[5] Math. Ann., 78, 1918, 405-415 = *Ges. Abh.*, 3, 145-156.

ideales de los objetos físicos o de los sucesos. Incluso conceptos tales como funciones o derivadas estaban implicados en los fenómenos reales y servían para describirlos.

Aparte de lo que ya hemos señalado como soporte de esta opinión acerca de las matemáticas, las afirmaciones de los matemáticos sobre el número de dimensiones que cabe considerar en geometría muestran claramente lo ligada que se mantenía la matemática a la realidad. Así, en el primer libro de *El Cielo*, Aristóteles dice: «La línea tiene una magnitud, el plano dos, y el sólido tres, y más allá de esas tres no hay ninguna otra magnitud, porque ellas lo llenan todo... No se puede pasar a otra cosa, como se pasa de la longitud al área y del área al volumen.» Y en otro párrafo: «... No hay magnitud que trascienda al tres, porque no hay más que tres dimensiones», y añade: «tres es el número perfecto». En su *Algebra*, John Wallis consideraba a un espacio con más de tres dimensiones como «un monstruo de la naturaleza, menos posible que una quimera o un centauro». Y sigue: «Longitud, Anchura y Espesor llenan todo el espacio. Tampoco Fansie puede imaginar cómo podría hacer una Cuarta Dimensión Local más allá de esas Tres.» Cardano, Descartes, Pascal y Leibniz también consideraron la posibilidad de una cuarta dimensión, rechazándola como absurda. Mientras el álgebra se mantuvo ligada a la geometría también se rechazó el producto de más de tres cantidades. Jacques Ozanam señalaba que el producto de más de tres letras representaría una magnitud «con tantas dimensiones como letras haya, pero sería sólo imaginaria, porque en la naturaleza no conocemos ninguna cantidad que tenga más de tres dimensiones».

La idea de una geometría matemática de más de tres dimensiones se rechazó incluso a comienzos del siglo XIX. Möbius, en su *Der barycentrische Calcul* (1827), señalaba que las figuras geométricas que no podían superponerse en tres dimensiones al ser cada una de ellas imagen especular de la otra, sí podrían superponerse en cuatro dimensiones. Pero añadía a continuación:[6] «Sin embargo, como tal espacio no puede ser pensado, la superposición es imposible.» Kummer, hacia 1860, todavía se burlaba de la idea de una geometría tetradimensional. Las objeciones que todos esos autores planteaban frente a una geometría de dimensión más elevada eran razonables en tanto que se identificaba la geometría con el estudio del espacio físico.

Pero gradual e inconscientemente los matemáticos comenzaron a

---

[6] *Ges. Werke*, 1, 172.

introducir conceptos que tenían poco o ningún significado físico directo. Entre ellos, los más turbadores eran los números negativos y complejos. Como esos dos tipos de números no existían «realmente» en la naturaleza, eran todavía sospechosos a comienzos del siglo XIX, aunque ya se utilizaban ampliamente entonces. La representación geométrica de los números negativos como distancias en una dirección sobre una recta y la de los números complejos como puntos o vectores en el plano complejo que, como Gauss subrayó acerca de los últimos, les dio significado intuitivo haciéndolos así admisibles, pudo también retrasar la conciencia de que la matemática trata con conceptos ideados por el hombre. Pero la introducción de los cuaterniones, las geometrías no euclídeas, los elementos complejos en geometría, la geometría n-dimensional, las funciones extrañas y los números transfinitos obligaron a reconocer la artificialidad de las matemáticas.

En relación con esto ya hemos indicado el impacto de las geometrías no euclídeas (cap. 36, sec. 8), y nos detendremos ahora en el efecto producido por la geometría n-dimensional. El concepto aparece ya, de forma inocua, en los trabajos analíticos de d'Alembert, Euler y Lagrange. D'Alembert sugirió la posibilidad de considerar al tiempo como una cuarta dimensión en su artículo de la *Encyclopédie* sobre «Dimension». Lagrange, al estudiar la reducción de formas cuadráticas a su expresión canónica, introdujo informalmente formas en n variables, y utilizó también el tiempo como una cuarta dimensión en su *Mécanique Analitique* (1788) y en su *Théorie des fonctions analytiques* (1797). En esta última obra, decía: «Podemos así considerar a la mecánica como una geometría de cuatro dimensiones, y a la mecánica analítica como una extensión de la geometría analítica.» Los trabajos de Lagrange ponían así a un mismo nivel las tres coordenadas espaciales y una cuarta para representar el tiempo. Más tarde, George Green ya no vacilaba en su artículo de 1828 sobre teoría del potencial en considerar problemas de potencial en n dimensiones, diciendo de esa teoría: «Ya no está confinada, como antes, a las tres dimensiones del espacio.»

Esos primeros desarrollos en n dimensiones no se planteaban como una investigación propiamente geométrica; eran generalizaciones naturales de trabajos analíticos que ya no estaban ligados a la geometría. En parte, esa introducción del lenguaje n-dimensional se planteaba únicamente como algo conveniente, que ayudaba al pensamiento analítico. Era útil pensar en $(x_1,x_2, ..., x_n)$ como un punto,

y en una ecuación en *n* variables como una hipersuperficie en el espacio *n*-dimensional, porque pensando en términos de lo que eso significa en la geometría tridimensional se podía introducir cierta intuición en el trabajo analítico. De hecho, Cauchy [7] subrayó que el concepto de espacio *n*-dimensional es útil en muchas investigaciones analíticas, especialmente las de teoría de números.

No obstante, el estudio serio de la geometría *n*-dimensional, sin que implicara un espacio físico de *n* dimensiones, se emprendió también en el siglo XIX; el fundador de esta geometría abstracta fue Grassmann, en su *Ausdehnungslehre* de 1844. Ahí se encuentra el concepto de geometría *n*-dimensional en su completa generalidad. En una nota publicada en 1845 decía Grassmann:

> Mi *Ausdehnungslehre* construye el fundamento abstracto de la teoría del espacio; esto es, queda libre de toda intuición espacial y es una ciencia puramente matemática; la geometría constituye únicamente su aplicación especial al espacio (físico).
>
> Sin embargo, los teoremas que presento en él no son simples traducciones a un lenguaje abstracto de resultados geométricos; tienen una significación mucho más general, porque mientras que la geometría ordinaria se limita a las tres dimensiones del espacio (físico), la ciencia abstracta queda libre de esa limitación.

Grassmann añade que la geometría ordinaria es considerada impropiamente como una rama de la matemática pura, ya que se trata en realidad de una rama de la matemática aplicada que trata de un tema no creado por el intelecto sino exterior a él: la materia. Pero, dice, es posible crear una disciplina puramente intelectual que se ocupe de la extensión como concepto, más que del espacio percibido por las sensaciones. Así, la obra de Grassmann aparece como representativa de la corriente que afirma que el pensamiento puro puede producir construcciones arbitrarias, que pueden ser o no físicamente aplicables.

Cayley, independientemente de Grassmann, se propuso también tratar analíticamente la geometría *n*-dimensional «sin recurrir a nociones metafísicas». En el *Cambridge Mathematical Journal* de 1845 publicó unos «Capítulos de Geometría Analítica en *N* Dimensiones»,[8] en los que ofrecía resultados analíticos en *n* variables, que

---

[7] Comp. Rend., 24, 1847, 885-887 = *Œuvres*, (1), 10, 292-295.

[8] 4. 119-127 = *Collected Math. Papers*, 1, 55-62.

para $n = 3$ se concretaban en teoremas ya conocidos sobre superficies. Aunque no hizo nada especialmente nuevo en geometría n-dimensional, el concepto de ésta quedaba plenamente establecido en esos trabajos.

Cuando Riemann presentó su *Habilitationsvortrag* en 1854, «Uber die Hypotesen welche die Geometrie zu Grunde liegen», no vaciló en tratar variedades n-dimensionales, aunque su objetivo principal era la geometría del espacio físico tridimensional. Los que siguieron el camino abierto por ese trabajo básico —Helmholtz, Lie, Christoffel, Beltrami, Lipschitz, Darboux y otros— continuaron trabajando en el espacio n-dimensional.

La noción de geometría n-dimensional encontró tenaz resistencia en algunos matemáticos incluso mucho tiempo después de haber sido introducida. Aquí, como en el caso de los números negativos y complejos, la matemática progresaba más allá de los conceptos sugeridos por la experiencia, y los matemáticos todavía tenían que asumir que su objeto podía consistir en conceptos creados por la mente, y que ya no era, si es que alguna vez lo había sido, una lectura de la naturaleza.

No obstante, en torno a 1850, fue ganando aceptación la idea de que la matemática puede introducir y tratar conceptos y teorías bastante arbitrarios que no tienen una interpretación física inmediata, pero que sin embargo pueden ser útiles, como en el caso de los cuaterniones, o satisfacer un deseo de generalidad, como en el caso de la geometría n-dimensional. Hankel, en su *Theorie der complexen Zahlensysteme* (1867, p. 10) defendió la matemática como «puramente intelectual, una teoría pura de las formas, que tiene como objeto no la combinación de cantidades o de sus imágenes, los números, sino asuntos del pensamiento a los que pueden corresponder objetos o relaciones reales, si bien esa correspondencia no es precisa».

Cantor, en defensa de su creación de los números transfinitos como cantidades realmente existentes, aducía que la matemática se distingue de otras ciencias por su libertad para crear sus propios conceptos sin atender a la realidad transitoria. En 1883 [9] escribía: «La matemática es completamente libre en su desarrollo, y sus conceptos sólo se ven restringidos por la necesidad de ser no contradictorios y estar coordinados con los conceptos previamente introducidos mediante definiciones precisas... La esencia de la matemática

---

[9] Math. Ann., 21, 1883, 563-564 = *Ges. Abh.*, 182.

reside en su libertad.» Prefería el término «matemática libre» al más usual de «matemática pura».

Esta nueva visión de las matemáticas se extendió a sus ramas más antiguas y físicamente asentadas. En su *Universal Algebra* (1898), Alfred North Whitehead decía (p. 11):

> ... el álgebra no depende de la aritmética para la validez de sus leyes de transformación. Si hubiera tal dependencia es obvio que tan pronto como se volvieran aritméticamente ininteligibles las expresiones algebraicas, todas las leyes referidas a ellas tendrían que perder su validez. Pero las leyes del álgebra, aunque sugeridas por la aritmética, no dependen de ella; dependen por entero de convenios mediante los que se establece que ciertos modos de agrupar los símbolos deben considerarse como idénticos, asignándose así ciertas propiedades a los signos que forman los símbolos del álgebra.

El álgebra es un desarrollo lógico independiente del significado. «Es obvio que podemos tomar los signos que queramos y manipularlos de acuerdo con cualquier regla que decidamos asignarles» (p. 4). Whitehead indica que tales manipulaciones arbitrarias pueden ser frívolas, y que sólo son significativas las construcciones a las que se puede atribuir algún sentido, o que pueden utilizarse de alguna forma.

También la geometría cortó sus amarras con la realidad física. Como señalaba Hilbert en sus *Grundlagen* de 1899, la geometría habla de cosas cuyas propiedades quedan especificadas en los axiomas. Aunque Hilbert se refería únicamente a la estrategia con la que se deben enfocar las matemáticas con objeto de examinar su estructura lógica, apoyó y alentó sin embargo la opinión de que la matemática es algo completamente distinto de los conceptos y leyes de la naturaleza.

4. La pérdida de la verdad

La introducción y la aceptación gradual de conceptos que no tienen contrapartida inmediata en el mundo real obligó naturalmente a reconocer que la matemática es una creación humana, en cierto modo arbitraria, más que una idealización de las realidades de la naturaleza, obtenida únicamente a partir de ésta. Pero acompañando a este reconocimiento, y empujando de hecho a su aceptación, se produjo un descubrimiento más profundo: que la matemática no es un conjunto de verdades acerca de la naturaleza. El desarrollo que

planteó la cuestión de la verdad fue la geometría no euclídea, aunque su impacto se vio aplazado por el característico conservadurismo y estrechez de miras de la mayoría de los matemáticos. El filósofo David Hume (1711-1776) había señalado ya que la naturaleza no se acomoda a patrones fijos ni leyes necesarias; pero la opinión dominante, expresada por Kant, era que las propiedades del espacio físico eran euclídeas. Incluso Legendre, en sus *Eléments de géométrie* de 1794, creía todavía que los axiomas de Euclides eran verdades autoevidentes.

Con respecto al menos a la geometría, la opinión que hoy parece correcta fue expresada primeramente por Gauss. A comienzos del siglo XIX ya estaba convencido de que la geometría es una ciencia empírica, que debe alinearse junto a la mecánica, mientras que la aritmética y el análisis eran para él verdades a priori. En 1830 escribía a Bessel:[10]

> Según mi convicción más profunda, la teoría del espacio ocupa un lugar en nuestro conocimiento a priori completamente diferente del que ocupa la aritmética pura. En todo nuestro conocimiento de la primera falta la convicción completa de necesidad (también de verdad absoluta) que caracteriza a la segunda; debemos añadir humildemente que si el número es un mero producto de nuestra mente, el espacio tiene una realidad fuera de ella cuyas leyes no podemos prescribir completamente a priori.

Sin embargo parece que Gauss se hallaba en cierta contradicción consigo mismo, porque también expresó la opinión de que toda la matemática es creación humana. En otra carta a Bessel, del 21 de noviembre de 1811, en la que hablaba de funciones de una variable compleja, decía:[11] «No se debe olvidar nunca que las funciones, como cualquier otra construcción matemática, son sólo nuestras propias creaciones, y que cuando la definición con la que uno comienza deja de tener sentido, no debería preguntarse qué es, sino qué tendría que suponer para que siguiera siendo significativa.»

Pese a las opiniones de Gauss acerca de la geometría, la mayoría de los matemáticos pensaban que en ella había verdades básicas. Bolyai creía que las verdades absolutas de la geometría eran los axiomas y teoremas comunes a la euclídea y la hiperbólica. No conocía

---

[10] *Werke*, 8, 201.
[11] *Werke*, 10, 363.

la geometría elíptica, y en su época todavía no se podía concebir que muchos de esos axiomas comunes no lo eran para todas las geometrías.

En su artículo de 1854 «Sobre las hipótesis en que se basa la Geometría», Riemann creía todavía que había ciertas proposiciones acerca del espacio que serían a priori, aunque no estuviera incluida entre ellas la afirmación de que el espacio físico es verdaderamente euclídeo. Era, sin embargo, localmente euclídeo.

Cayley y Klein seguían apegados a la realidad de la geometría euclídea (vid. cap. 38, sec. 6). En su discurso presidencial a la Asociación Británica para el Avance de la Ciencia,[12] decía: «... no que las proposiciones de la geometría sean sólo aproximadamente ciertas, sino que permanecen absolutamente ciertas con respecto a ese espacio euclídeo que ha sido considerado durante tanto tiempo como el espacio físico de nuestra experiencia.» Aunque ambos había trabajado en geometrías no euclídeas, las consideraban como novedades resultantes de introducir otras distancias en la geometría euclídea, y no eran capaces de ver que las geometrías no euclídeas son tan básicas y tan aplicables como la euclídea.

En los años noventa Bertrand Russell se planteó la cuestión de qué propiedades del espacio son necesarias para la experiencia y están implícitas en ésta. Es decir, aquellas propiedades a priori cuya negación convertiría en un sinsentido la experiencia. En su *Essay on the Foundations of Geometry* (1897) acepta que la geometría euclídea no representa un conocimiento a priori, pero defiende que la geometría proyectiva es a priori para cualquier planteamiento geométrico, conclusión que podemos entender a la luz de la importancia de ese enfoque en torno a 1900. A continuación añade como a priori para la geometría los axiomas comunes a la euclídea y todas las no euclídeas. La homogeneidad del espacio, la dimensionalidad finita y algún concepto de distancia hacen posibles las mediciones. Considera sin embargo como empíricos los hechos de que el espacio sea tridimensional y euclídeo.

Russell considera como un resultado técnico sin significación filosófica el que se puedan derivar las geometrías métricas de la proyectiva introduciendo una distancia. La geometría métrica, según él, es una rama de la matemática separada y lógicamente subsidiaria, y

[12] Report of the Brit. Assn. for the Adv. of Sci., 1883, 3-37 = *Coll. Math. Papers*, 11, 429-459.

no es a priori. Con respecto a las geometrías euclídea y no euclídeas, se distancia de Cayley y Klein considerándolas a todas ellas igualmente significativas. Como los únicos espacios métricos que poseen las propiedades indicadas más arriba son los euclídeos, hiperbólicos y simple o doblemente elípticos, concluye que esas son las únicas geometrías métricas posibles, y desde luego la euclídea es la única físicamente aplicable; las otras tienen importancia filosófica, mostrando que puede haber otras geometrías. Retrospectivamente podemos ahora decir que Russell sustituía el prejuicio euclídeo por un prejuicio proyectivo.

Aunque los matemáticos fueron lentos en reconocer el hecho, claramente observado por Gauss, de que no hay ninguna seguridad en la verdad física de la geometría euclídea, gradualmente fueron llegando a esa convicción y también a la sostenida por Gauss de que la verdad de la matemática reside en la aritmética y consecuentemente también en el análisis. Kronecker, por ejemplo, en su ensayo «Uber den Zahlbegriff» (Sobre el Concepto de Número),[13] sostuvo la verdad de las disciplinas aritméticas, negándosela a la geometría. Gottlob Frege, acerca de cuya obra hablaremos más tarde, también insistió en la verdad de la aritmética.

Sin embargo, también la aritmética y el análisis construido sobre ella se hicieron pronto sospechosos. La creación de álgebras no conmutativas, en particular las de cuaterniones y matrices, plantearon naturalmente la cuestión de cómo podría estar uno seguro de que los números ordinarios poseyeran la propiedad privilegiada de la verdad acerca del mundo real. El ataque contra la veracidad de la aritmética vino en primer lugar de Helmholtz. Después de haber insistido, en un ensayo famoso,[14] en que nuestro conocimiento del espacio físico proviene únicamente de la experiencia y depende de la existencia de cuerpos rígidos que puedan servir, entre otras cosas, como reglas para medir, en su *Zählen und Messen* (Contar y Medir, 1887) puso en cuestión las verdades de la aritmética. Consideraba como el principal problema de ésta el significado o la validez de la aplicación objetiva de la cantidad y la igualdad a la experiencia. La aritmética misma puede considerarse como una exposición consistente de las consecuencias de las operaciones aritméticas. Trata con

[13] Jour. für Math., 101, 1887, 337-353 = *Werke*, 3, 249-274.

[14] Nachrichten Konig. Ges. der Wiss. zu Gilt., 15, 1868, 193-221 = Wiss. Abh., 2, 618-639.

símbolos y puede entenderse como un juego. Pero esos símbolos se aplican a objetos reales y a las relaciones entre ellos, y proporcionan resultados acerca del funcionamiento real de la naturaleza. ¿Cómo es esto posible? ¿Bajo qué condiciones se pueden aplicar a los objetos reales los números y las operaciones aritméticas? En particular, ¿cuál es el significado objetivo de la igualdad de dos objetos, y cómo puede tratarse la adición física como suma aritmética?

Helmholtz señalaba que la aplicabilidad de los números no es ni un accidente ni tampoco una prueba de la verdad de las leyes numéricas. *Algunos* tipos de experiencia los sugieren, y a esos es a los que se pueden aplicar. Para aplicar números a objetos reales, decía Helmholtz, éstos no deben desaparecer, o mezclarse entre sí, o dividirse. Al añadir físicamente una gota de lluvia a otra no se obtienen dos gotas de lluvia. Sólo la experiencia puede decirnos si los objetos de una colección física determinada mantienen en ella su identidad de manera que se pueda atribuir a la colección un número definido de objetos. De igual modo, saber cuándo se puede aplicar la igualdad entre dos cantidades físicas también depende de la experiencia. Cualquier afirmación de igualdad cuantitativa debe satisfacer dos condiciones: si se intercambian los objetos, deben seguir siendo iguales. Y si el objeto *a* es igual a *c* y el objeto *b* es igual a *c*, los objetos *a* y *b* deben ser iguales. De esta manera podemos hablar de la igualdad de pesos o intervalos de tiempo, ya que para esos objetos sí se puede determinar la igualdad. Pero dos sonidos pueden ser indistinguibles para el oído de un tercero, intermedio entre ambos, aunque sí puedan distinguirse entre sí. Aquí dos cosas iguales a una tercera no son iguales entre sí. Tampoco se pueden sumar los valores de unas resistencias eléctricas conectadas en paralelo para obtener la resistencia total, ni se pueden combinar de cualquier manera los índices de refracción de diferentes medios.

Hacia finales del siglo XIX prevaleció la opinión de que todos los axiomas de las matemáticas son arbitrarios, constituyendo únicamente la base para deducir consecuencias de ellos. Como ya no se entendían como verdades acerca de los conceptos implicados en ellos, dejó de importar el significado físico de esos conceptos. Ese significado podía, a lo más, servir como guía heurística cuando los axiomas mantenían alguna relación con la realidad. Así, incluso los conceptos quedaron separados del mundo físico. Hacia 1900 las matemáticas se habían despegado de la realidad, abandonando clara e irreparablemente su pretensión a la verdad acerca de la naturaleza,

y convirtiéndose en la búsqueda de las consecuencias necesarias de axiomas arbitrarios acerca de cosas sin sentido.

La pérdida de la verdad y la aparente arbitrariedad, la naturaleza subjetiva de las ideas y resultados matemáticos, turbaron profundamente a muchos, que entendieron esto como una denigración de la matemática. Algunos de ellos adoptaron una postura mística que de alguna manera garantizara cierta realidad y objetividad a las matemáticas. Esos matemáticos suscribieron la idea de que la matemática es una realidad en sí misma, un cuerpo independiente de verdades, y que sus objetos nos vienen dados como nos vienen dados los objetos del mundo real; los matemáticos simplemente descubren los conceptos y sus propiedades. Hermite, por ejemplo, en una carta a Stieltjes,[15] decía: «Creo que los números y las funciones del análisis no son el producto arbitrario de nuestras mentes; creo que existen fuera de nosotros con el mismo carácter de necesidad que los objetos de la realidad objetiva, y que nosotros los encontramos o descubrimos y los estudiamos como lo hacen los físicos, químicos y zoólogos.»

En el Congreso Internacional de Bolonia de 1928, Hilbert decía:[16] «¿Qué sucedería con la verdad de nuestro conocimiento y con la existencia y el progreso de la ciencia si no hubiese ninguna verdad en las matemáticas? De hecho, actualmente aparece con demasiada frecuencia en los escritos profesionales y en lecciones orales cierto escepticismo o falta de confianza acerca del conocimiento; se trata de un tipo de ocultismo que considero perjudicial.»

Godfrey H. Hardy (1877-1947), sobresaliente analista del siglo XX, decía en 1928:[17] «Los teoremas matemáticos son verdaderos o falsos; su verdad o falsedad es absolutamente independiente de nuestro conocimiento de ellos. En *cierto* sentido, la verdad matemática forma parte de la realidad objetiva.» Expresó de nuevo la misma opinión en su libro *A Mathematician's Apology* (ed. de 1967, p. 123): «Creo que la realidad matemática está fuera de nosotros, que nuestra función consiste en descubrirla u observarla, y que los teoremas que describimos con grandilocuencia como nuestras «creaciones» son simplemente las anotaciones de nuestra observación.»

---

[15] *C. Hermite-T. Stieltjes correspondance,* Gauthier-Villars, 1905, 2, 398.
[16] Atti del Congreso, 1, 1929, 141 = *Grundlagen der Geometrie,* 7.ª ed., 323.
[17] Mind, 38, 1929, 1-25.

5. La matemática como el estudio de estructuras arbitrarias

Los matemáticos del siglo XIX se ocupaban prioritariamente del estudio de la naturaleza, y la física fue sin duda la fuente de inspiración más importante para el trabajo matemático. Los investigadores más descollantes —Gauss, Riemann, Fourier, Hamilton, Jacobi, Poincaré— y otros menos conocidos —Christoffel, Lipschitz, Du Bois-Reymond, Beltrami...— trabajaron directamente en problemas de física y en problemas matemáticos surgidos de investigaciones físicas. Incluso los autores considerados comúnmente como matemáticos puros, como por ejemplo Weierstrass, trabajaron en problemas físicos. De hecho, los problemas físicos proporcionaron más sugerencias e indicaciones para las investigaciones matemáticas que en cualquier otro siglo anterior, y para analizarlos se creó una matemática altamente compleja. Fresnel había señalado que «la naturaleza no se detiene ante dificultades de análisis», pero los matemáticos no se arredraron y consiguieron vencer esas dificultades. La única rama importante que se había mantenido por una satisfacción intrínsecamente estética, al menos desde la obra de Diofanto, era la teoría de números.

No obstante, fue por primera vez en el siglö XIX cuando los matemáticos no sólo desarrollaron su trabajo más allá de las necesidades de la ciencia y la tecnología de su época, sino que plantearon y resolvieron cuestiones que no tenían que ver con problemas reales. La *raison d'être* de este desarrollo se podría describir así: la convicción bimilenaria de que la matemática constituía la verdad acerca de la naturaleza se había hecho pedazos. Pero las teorías matemáticas que ahora se reconocían como arbitrarias habían demostrado sin embargo su utilidad en el estudio de la naturaleza. Aunque las teorías existentes debían mucho históricamente a sugerencias de la naturaleza, quizá nuevas teorías construidas únicamente por la mente podrían ser también útiles para su representación. Los matemáticos se sintieron entonces libres para crear estructuras arbitrarias, quedando justificada una nueva libertad en la investigación matemática. Sin embargo, como algunas de las estructuras ya creadas hacia 1900, y muchas de las que vendrían después, parecían tan artificiales y alejadas de cualquier aplicación potencial, sus patrocinadores comenzaron a defenderlas como deseables en y por sí mismas.

El ascenso y aceptación gradual de la opinión de que la matemática debía ocuparse de estructuras arbitrarias que no estaban obli-

gadas a tener nada que ver, en primera ni última instancia, con el estudio de la naturaleza, condujo a un cisma que se conoce actualmente como el de la matemática pura frente a la aplicada. Semejante ruptura con la tradición no podía dejar de generar controversia. Nos limitaremos únicamente a citar unos pocos de los argumentos de cada bando.

Fourier había escrito, en el prefacio a su *Théorie Analytique de la Chaleur:* «El estudio profundo de la naturaleza es el campo más fértil para los descubrimientos matemáticos. Ese estudio ofrece no sólo la ventaja de un objetivo bien definido, sino también la de excluir cuestiones vagas y cálculos inútiles. Es un medio para construir el análisis en sí mismo y para descubrir qué ideas importan verdaderamente y cuáles debe preservar la ciencia. Las ideas fundamentales son aquellas que representan los acontecimientos naturales.» También subrayó la aplicación de las matemáticas a problemas socialmente útiles.

Aunque Jacobi había hecho investigaciones de primer orden en mecánica y astronomía, criticó vivamente las opiniones de Fourier. El 2 de julio de 1830 escribía a Legendre:[18] «Es cierto que Fourier piensa que el objeto prioritario de la matemática es la utilidad pública y la explicación de los fenómenos naturales; pero un científico como él debería saber que el único objeto de la ciencia es el honor del espíritu humano y sobre esta base una cuestión de (teoría de) números es tan importante como una cuestión acerca del sistema planetario...»

A lo largo del siglo, la deriva de más y más investigadores hacia la matemática pura suscitó reacciones de protesta. Kronecker, por ejemplo, escribía a Helmholtz: «La riqueza de su experiencia práctica con problemas sanos e interesantes dará un nuevo sentido y un nuevo ímpetu a la matemática... La especulación matemática unilateral e introspectiva conduce a campos estériles.»

Felix Klein, en su *Teoría Matemática del Giróscopo* (1897, pp. 1-2) afirmaba: «La gran necesidad del presente en la ciencia matemática es que la ciencia pura y los departamentos de ciencias físicas en los que encuentra sus más importantes aplicaciones vuelvan de nuevo a la asociación íntima que se mostró tan fructífera en las obras de Lagrange y Gauss.» Y Emile Picard, a comienzos de este siglo

---

[18] *Ges. Werke*, 1, 454-455.

(*La Science moderne et son état actuel*, 1908) prevenía contra la tendencia a las abstracciones y los problemas sin interés.

Un poco después, Felix Klein insistía de nuevo sobre el asunto.[19] Temiendo el abuso de la libertad para crear estructuras arbitrarias, subrayaba que éstas son «la muerte de la ciencia. Los axiomas de la geometría no son... proposiciones arbitrarias, sino que se deducen, en general, de nuestra percepción espacial, y están determinados en cuanto a su contenido preciso por la utilidad.» Para justificar los axiomas no euclídeos, Klein señalaba que la visualización puede verificar el axioma euclídeo de las paralelas sólo hasta ciertos límites. En otra ocasión indicaba que «quienquiera que reivindique el privilegio de la libertad debe también soportar la responsabilidad». Klein entendía por «responsabilidad» la dedicación a investigar la naturaleza.

A pesar de las advertencias, la tendencia a la abstracción, a generalizar porque sí los resultados existentes, y la búsqueda de problemas arbitrariamente planteados, prosiguió en nuestro siglo. La razonable necesidad de estudiar toda una clase de problemas para saber más acerca de algún caso concreto, y de abstraer para quedarse con lo esencial de un problema, se convirtieron en excusas para complacerse en generalidades y abstracciones en y por sí mismas.

En parte para contrarrestar esa tendencia a la generalización, Hilbert no sólo insistió en que los problemas concretos constituyen el flujo vital de las matemáticas, sino que se tomó el trabajo de publicar en 1900 una lista de los 23 más sobresalientes (vid. bibliografía) y de citarlos en una conferencia que dio en el Segundo Congreso Internacional de Matemáticos en París. El prestigio de Hilbert llevó a muchos a ocuparse de esos problemas. Ningún honor podía sobrepasar al de resolver un problema planteado por alguien tan respetado. Pero la tendencia a la creación libre y a las abstracciones y generalizaciones no pudo contenerse. La matemática se despegó de la naturaleza y la ciencia para seguir su propio curso.

---

[19] *Elementary Mathematics from an advanced standpoint*, Mac Millan, 1939; Dover (reimp.), 1945, vol. 2, 187.

6. El problema de la consistencia

La matemática, desde el punto de vista de la lógica, era a finales del siglo XIX una colección de estructuras construida cada una de ellas sobre su propio sistema de axiomas. Como ya hemos señalado, una de las propiedades necesarias de cualquier estructura de ese tipo es la consistencia de sus axiomas. Mientras se consideró la matemática como la verdad acerca de la naturaleza, no cabía la posibilidad de que pudieran surgir teoremas contradictorios, y pensarlo siquiera habría sido considerado como absurdo. Cuando se crearon las geometrías no euclídeas, su aparente discrepancia con la realidad planteó la cuestión de su consistencia. Como hemos visto, se respondió a esta cuestión haciendo depender la consistencia de las geometrías no euclídeas de la de la geometría euclídea.

En los años ochenta, la conciencia de que ni la aritmética ni la geometría euclídea son verdades en sí hizo imperativa la investigación de la consistencia de esas ramas de la matemática. Peano y su escuela comenzaron a considerar este problema en los años noventa. Creía que se podrían concebir pruebas claras que dilucidaran la cuestión, pero los acontecimientos demostraron que estaba equivocado. Hilbert consiguió establecer la consistencia de la geometría euclídea sobre la hipótesis de que la aritmética es consistente (cap. 42, sec. 3). Pero la consistencia de esta última no había quedado establecida, y Hilbert planteó este problema como el segundo de su lista en el Segundo Congreso Internacional de 1900; en su «Axiomatisches Denken»,[20] lo señaló como el problema básico de los fundamentos de la matemática. Muchos otros investigadores eran también conscientes de la importancia del problema. En 1904 Alfred Pringsheim (1850-1941),[21] afirmó que la verdad que persigue la matemática no es ni más ni menos que la consistencia. En el capítulo 51 examinaremos los trabajos sobre este problema.

7. Una mirada hacia el futuro

El ritmo de la creación matemática se ha venido incrementando sin pausa desde 1600, y esto sigue siendo cierto para el siglo XX, en

[20] Math. Ann. 78, 1918, 405-415 = *Ges. Abb.*, 145-156.
[21] Jahres. der Deut. Math.-Verein, 13, 1904, 381.

el que se desarrollaron la mayoría de los campos explorados en el XIX. No obstante, los detalles de los nuevos trabajos en esos campos interesarían únicamente a los especialistas. Limitaremos por tanto nuestro examen del siglo XX a los campos que aparecieron o se hicieron por primera vez importantes en este período. Además, nos ocuparemos tan sólo de los comienzos de esas investigaciones. Los desarrollos del segundo y tercer cuartos de este siglo son todavía demasiado recientes como para poderlos evaluar correctamente. Ya hemos visto como muchas áreas exploradas vigorosa y entusiásticamente en el pasado, consideradas por sus abogados como la esencia de la matemática, resultaron ser modas pasajeras o con poca influencia en el curso general de ésta. Por muy confiados que estén los matemáticos del último medio siglo en la gran importancia de sus trabajos, el lugar que ocupen sus contribuciones en la historia de la matemática no puede decidirse todavía.

Bibliografía

Fang, J.: *Hilbert*, Paideia Press, 1970. Esbozos de la obra matemática de Hilbert.

Hardy, G. H.: *A Mathematician's Apology*, Cambridge University Press, 1940 y 1967. Hay versión castellana, *Autojustificación de un matemático*, Barcelona, Ariel, 1981.

Helmholtz, H. von: *Counting and Measuring*, D. van Nostrand, 1930. Traducción al inglés de *Zählen und Messen*. *Wissenschaftliche Abhandlungen*, 4, 356-391.

— «Uber den Ursprung Sinn und Bedeutung der geometrischen Sätze»; traducción al inglés: «On the Origin and Significance of Geometrical Axioms», en Helmholtz: *Popular Scientific Lectures*, Dover (reimpresión), 1962, pp. 223-249. También en James R. Newman: *The World of Mathematics*, Simon & Schuster, 1956, vol. 1, pp. 647-668. Hay versión castellana, *Sigma, el mundo de las matemáticas*, Barcelona, Grijalbo, 8.ª ed., 1983. Véase también Helmholtz: *Wissenschaftliche Abhandlungen*, 2, 640-660.

Hilbert, David: «Sur les problèmes futurs des mathématiques». *Comptes Rendus du Deuxième Congrès International des Mathématiciens*, Gauthier-Villars, 1902, 58-114. También se puede encontrar en alemán en *Nachrichten König. Ges. der Wiss. zu Gött.*, 1900, 253-297, y en Hilbert: *Gesammelte Abhandlungen*, 3, 290-329. Hay una traducción al inglés en *Amer. Math. Soc. Bull.*, 8, 1901/1902, 437-479.

Klein, Felix: «Uber Arithmetisirung der Mathematik». *Ges. Math. Abh.*, 2, 232-240. Traducción al inglés en *Amer. Math. Soc. Bull.*, 2, 1895/1896, 241-249.
Pierpoint, James: «On the Arithmetization of Mathematics». *Amer. Math. Soc. Bull.*, 5, 1898/1899, 394-406.
Poincare, Henri: *The Foundations of Science*, Science Press, 1913. Vid. en particular pp. 43-91.
Reid, Constance: *Hilbert*, Springer-Verlag, 1970. Se trata de una biografía.

Capítulo 44

LA TEORIA DE FUNCIONES DE UNA O VARIAS VARIABLES REALES

> Si Newton y Leibniz hubieran llegado a imaginarse que las funciones continuas no tienen por qué tener necesariamente derivada (y esto es lo que ocurre, en general), nunca se habría creado el cálculo diferencial.
>
> EMILE PICARD

1. Los orígenes

La teoría de funciones de una o varias variables reales tuvo su origen en el intento de llegar a entender y clarificar cierto número de descubrimientos extraños que se habían ido haciendo a lo largo del siglo XIX. La aparición de funciones continuas pero no diferenciables, de series de funciones continuas cuya suma era discontinua, de funciones continuas que no eran monótonas a trozos, de funciones con derivadas acotadas pero que no eran integrables en el sentido de Riemann, de curvas rectificables pero que no lo eran de acuerdo con la definición de longitud de un arco de curva dada por el cálculo infinitesimal, y de funciones no integrables que eran límite de sucesiones de funciones integrables, todo ello, en fin, parecía contradecir flagrantemente el comportamiento que se esperaba de las funciones, las derivadas y las integrales. Otro tipo de motivaciones para estudiar más a fondo el comportamiento de las funciones provenía de las investigaciones sobre las series de Fourier. Esta teoría, tal como había sido construida por Dirichlet, Riemann, Cantor, Ulisse Dini (1845-1918), Jordan y otros matemáticos del siglo XIX, resultaba ser una herramienta muy satisfactoria para la matemática aplicada, pero

lo cierto era que las propiedades de estas series, en el estado de desarrollo alcanzado, no constituían aún una teoría que pudiera satisfacer al matemático puro. Todavía se echaban de menos las relaciones de unidad, simetría y completitud que debería haber entre funciones y series.

Las investigaciones que se desarrollaban en teoría de funciones hicieron énfasis en la teoría de la integral, porque parecía que la mayoría de las incongruencias se podrían resolver extendiendo este concepto, y por tanto estos trabajos pueden considerarse en gran medida como una continuación directa de la obra de Riemann, Darboux, Du Bois-Reymond, Cantor y otros (véase cap. 40, sec. 4).

2. La integral de Stieltjes

En realidad, la primera extensión del concepto de integral tuvo su origen en un tipo de problemas completamente distintos de los que acabamos de mencionar. En 1894 publicaba Thomas Jan Stieltjes (1856-1894) su largo artículo titulado «Recherches sur les fractions continues»,[1] trabajo de una gran originalidad en el que partía de un problema muy particular para terminar resolviéndolo con rara elegancia. Esta obra venía a sugerir problemas de un tipo completamente nuevo tanto en la teoría de funciones analíticas como en la de funciones de una variable real. En particular, y con objeto de representar el límite de una sucesión de funciones analíticas, se vio obligado Stieltjes a introducir una nueva integral que generalizase la idea de Riemann-Darboux.

Stieltjes comienza considerando una distribución positiva de masa a lo largo del eje Ox, lo que venía a generalizar el concepto de densidad puntual, que ya había sido utilizado, desde luego. Hace observar que tal distribución de masa vendrá dada por una función creciente $\phi(x)$ que exprese la masa total acumulada en el intervalo $[0,x]$ para cada $x > 0$, donde las discontinuidades de ϕ corresponderán a masas concentradas en un punto. Para tal distribución de masas sobre el intervalo $[a,b]$ define las sumas de Riemann

$$\sum_{i=0}^{n} f(\xi_i)(\phi(x_{i+1}) - \phi(x_i))$$

[1] Ann. Fac. Sci. de Toulouse, 8, 1894, J.1-122 y 9, A.1-47 = *Œuvres complètes*, 2, 402-559.

donde los puntos $x_0, x_1, \ldots, x_n$ forman una partición del intervalo $[a,b]$ y ξ_i está en $[x_i,x_{i+1}]$. Stieltjes demuestra después que cuando f es continua sobre $[a,b]$ y el subintervalo máximo de la partición tiende a cero, entonces la suma tiende a un límite que representa por $\int_a^b f(x)\, d\phi(x)$. A pesar de que utilizó esta integral en su propia obra, Stieltjes no investigó más sobre el concepto mismo de integral, excepto para definir de manera obvia, para el intervalo $(0,\infty)$

$$\int_0^\infty f(x)\, d\phi(x) = \lim_{b\to\infty} \int_0^b f(x)\, d\phi(x).$$

Este concepto de integral no fue adoptado por los matemáticos hasta mucho más tarde, cuando se le encontraron gran cantidad de aplicaciones (véase cap. 47, sec. 4).

3. Primeros trabajos sobre contenido y medida

Muy distinta fue la línea de pensamiento que condujo a otra generalización diferente de la idea de integral, la integral de Lebesgue. El estudio de los conjuntos de puntos de discontinuidad de funciones planteó el problema de cómo medir la extensión o «longitud» de tales conjuntos, porque precisamente la «extensión» de dichas discontinuidades es lo que determina la integrabilidad de la función. La teoría del contenido, y más tarde la teoría de la medida, fueron introducidas precisamente para extender la idea de longitud a conjuntos de puntos que no sean simples intervalos de la recta real usual.

El concepto de contenido se basa en la siguiente idea: considérese un conjunto E de puntos distribuidos de alguna manera sobre el intervalo $[a,b]$. En términos un poco imprecisos por el momento, supongamos que es posible encerrar o recubrir estos puntos mediante subintervalos pequeños de $[a,b]$ de manera que los puntos de E sean o bien interiores a uno de los intervalos o en el peor de los casos uno de sus extremos. Reduzcamos las longitudes de estos subintervalos cada vez más, añadiendo otros si es necesario para seguir recubriendo los puntos de E, a la vez que reducimos la suma de las longitudes de todos ellos. Al extremo inferior de las sumas de estos subintervalos que recubren los puntos de E se llama el contenido (exterior) de E. Esta formulación un poco vaga no es la del concepto

adoptado finalmente de manera definitiva, pero nos puede servir para entender lo que estaban intentando hacer los matemáticos.

Una definición de contenido (exterior) fue dada por Du Bois-Reymond en su obra *Die allgemeine Funktionentheorie* (1882), por Axel Harnack (1851-1888) en su *Die Elemente der Differential-und Integralrechnung* (1881), por Otto Stolz [2] y por Cantor.[3] Stolz y Cantor extendieron además la noción de contenido a conjuntos de dos y de más dimensiones utilizando rectángulos, paralelepípedos, etc. en lugar de intervalos.

La utilización de esta noción de contenido, que desgraciadamente no resultó ser satisfactoria en todos los sentidos, reveló sin embargo la existencia de conjuntos no densos en ninguna parte (es decir, contenidos en un intervalo pero no densos en ninguno de sus subintervalos) de contenido positivo, y que las funciones con tales conjuntos de puntos de discontinuidad no eran integrables en el sentido de Riemann. También había funciones con derivadas acotadas no integrables. Sin embargo, los matemáticos de la época, la década de los ochenta, seguían pensando que el concepto de integral de Riemann no se podría generalizar.

Con objeto de superar las limitaciones de la anterior teoría del contenido y de rigorizar la noción de área de una región plana, Peano introduce una definición de contenido más completa y muy mejorada en su obra *Applicazioni geometriche del calcolo infinitesimale* (1887). Peano introduce aquí las nociones de contenido interior y exterior de una región. Supongámonos situados en dos dimensiones; el contenido interior es el extremo superior de las áreas de todas las regiones poligonales contenidas en la región dada R, mientras que el contenido exterior es el extremo inferior de las áreas de todas las regiones poligonales que contengan a la región R. Si los contenidos interior y exterior coinciden, a esta valor común se le llama área de la región R. Para los conjuntos lineales la idea es análoga, utilizando esta vez intervalos en vez de polígonos. Peano hace observar que, si la función $f(x)$ es no negativa sobre el intervalo $[a,b]$, entonces

$$\underline{\int_a^b} f\,dx = C_i(R) \qquad \text{y} \qquad \overline{\int_a^b} = C_e(R),$$

[2] Math. Ann., 23, 1884, 152-156.

[3] Math. Ann., 23, 1884, 453-488 = *Ges. Abh.*, 210-246.

donde la primera integral es el extremo superior de las sumas de Riemann inferiores de f sobre $[a,b]$ y la segunda es el extremo inferior de las sumas de Riemann superiores, y $C_i(R)$ y $C_e(R)$ son los contenidos interior y exterior respectivamente de la región R limitada por el grafo de f y el eje Ox. Así pues, f es integrable si y sólo si R tiene área definida, en el sentido de que $C_i(R) = C_e(R)$.

A Jordan le corresponde el mérito de haber dado el paso más atrevido y definitivo en la teoría del contenido (*étendue*) de todo el siglo XIX. También introduce unos contenidos interior y exterior.[4] pero formula estos conceptos de una manera un poco más eficaz. Su definición para un conjunto de puntos E contenido en un intervalo $[a,b]$ comienza por el contenido exterior; hay que recubrir E por un conjunto finito de subintervalos de $[a,b]$ tales que cada punto de E sea interior o extremo de uno de esos subintervalos. Al extremo inferior de las sumas de las longitudes de todos los conjuntos de ese tipo de subintervalos que contienen al menos un punto de E se le llama contenido exterior de E. El contenido interior de E se define como el extremo superior de las sumas de los subintervalos de $[a,b]$ que contengan sólo puntos de E. Si los contenidos interior y exterior de E son iguales, se dice que E tiene contenido definido. Jordan aplicó la misma idea a conjuntos de un espacio n-dimensional cualquiera, reemplazando los subintervalos por rectángulos y sus análogos en dimensiones mayores. Ahora ya podía demostrar Jordan lo que llamó propiedad de aditividad: el contenido de la suma de un número *finito* de conjuntos disjuntos con contenido definido es la suma de los contenidos de los mismos. Esto no se verificaba en las teorías del contenido anteriores, excepto en la de Peano.

El interés de Jordan por el contenido provenía del intento de precisar la teoría de las integrales dobles extendidas a una región plana E. La definición generalmente adoptada consistía en dividir el plano en cuadrados R_{ij} por medio de rectas paralelas a los ejes de coordenadas. Esta partición del plano induce una partición de E en regiones E_{ij}. Entonces, por definición,

$$\int_E f(x, y)\, dE = \lim_{\Delta x, \Delta y \to 0} \sum_{i,j} f(x_i, y_j) a(R_{ij}),$$

donde $a(R_{ij})$ representa el área de R_{ij}, y la suma se extiende a todas

[4] Jour. de Math., 8, 1892, 69-99 = *Œuvres*, 4, 427-457.

las R_{ij} totalmente contenidas en E y a todas las R_{ij} que contengan puntos de E así como puntos exteriores a E. Para que exista la integral es necesario demostrar que la contribución de los R_{ij} que no están completamente contenidos en E puede despreciarse, es decir, que la suma de las áreas de los R_{ij} que contienen puntos frontera de E tiende a cero con las dimensiones de los R_{ij}. En general, hasta entonces se había supuesto que éste era el caso, y Jordan mismo lo hizo así en la primera edición de su *Cours d'Analyse* (vol. 2, 1883). Sin embargo, el descubrimiento de curvas tan peculiares como la de Peano, que llena un cuadrado, forzó a los matemáticos a ser más cautelosos. Si E tiene contenido de Jordan bidimensional bien definido, entonces se pueden despreciar los R_{ij} que contienen la frontera de E. Jordan también consiguió demostrar resultados sobre el cálculo de integrales dobles por integración iterada.

La segunda edición del *Cours d'Analyse* (vol. 1, 1893) de Jordan incluye ya su tratamiento del contenido y su aplicación a la integración. Aunque superior a las de sus predecesores, la definición de contenido de Jordan no era del todo satisfactoria. Según ella, un conjunto abierto acotado podría no tener contenido definido, y el conjunto de los puntos racionales de un intervalo acotado no tenía contenido.

La etapa siguiente en la teoría del contenido se debe a Borel. Borel se vio conducido al estudio de la teoría de la medida, tal como él la llamó, trabajando con conjuntos de puntos en que convergen series que representan funciones de variable compleja. Su *Leçons sur la théorie des fonctions* (1898) contiene sus primeros trabajos importantes sobre el tema; Borel se dio cuenta de los defectos de las teorías anteriores del contenido y les puso remedio.

Cantor había demostrado ya que todo abierto U de la recta es unión de una familia *numerable* de intervalos abiertos disjuntos dos a dos. En vez de aproximar U encerrándolo en un conjunto finito de intervalos, Borel propone, utilizando el resultado de Cantor, definir la medida de un abierto acotado U como la suma de las longitudes de los intervalos componentes. A continuación define la medida de la suma de una cantidad *finita o numerable* de conjuntos medibles disjuntos como la suma de sus medidas individuales, y la medida del conjunto A-B, siendo A y B medibles y B *contenido en* A, como la correspondiente diferencia de las medidas. Con estas definiciones podía atribuir una medida a los conjuntos formados sumando cualquier cantidad finita o numerable de conjuntos medi-

bles disjuntos y a la diferencia de dos conjuntos medibles cualesquiera *A* y *B*, siempre que *A* contenga a *B*. A continuación estudia los conjuntos de medida 0 y demuestra que cualquier conjunto de medida mayor que 0 tiene que ser no numerable.

La teoría de la medida de Borel suponía una mejora indudable de las teorías del contenido de Peano y Jordan, pero no iba a ser aún la última palabra sobre el tema, ni estudió Borel sus aplicaciones a la teoría de integración.

4. La integral de Lebesgue

La generalización de la idea de medida e integral que se considera hoy como definitiva se debe a Henri Lebesgue (1875-1941), discípulo de Borel y más tarde profesor del Collége de France. Siguiendo las ideas de Borel y también las de Jordan y Peano, presentó por primera vez sus propias ideas sobre la medida y la integral en su tesis «Intégrale, longueur, aire».[5] Su obra vino a reemplazar todas las creaciones del siglo XIX y, en particular, mejoró la teoría de la medida de Borel.

La teoría de integración de Lebesgue se basa en su definición de medida de conjuntos de puntos, y ambas ideas se aplican a conjuntos del espacio *n*-dimensional. Como ejemplo, nos limitaremos al caso unidimensional. Sea *E* un conjunto de puntos contenido en el intervalo $[a,b]$. Los puntos de *E* pueden encerrarse como puntos interiores de una familia finita o *infinita numerable* de intervalos d_1, d_2, ... contenidos en $[a,b]$. (Los extremos de $[a,b]$ pueden considerarse, en su caso, como extremos de algún d_i.) Puede demostrarse que la familia de intervalos $\{d_i\}$ se puede sustituir por otra formada por intervalos δ_1, δ_2, ... no rampantes tales que todo punto de *E* o es punto interior de uno de estos intervalos o el extremo común de dos intervalos adyacentes. Sea $\Sigma\ \delta_n$ la suma de las longitudes de los δ_i. El extremo inferior de las $\Sigma\ \delta_n$ para todas las posibles familias $\{\delta_i\}$ recibe el nombre de medida exterior de *E* y se representa por $m_e(E)$. La medida interior de *E*, $m_i(E)$, viene definida como la medida exterior del conjunto $C(E)$, es decir, del complemento de *E* con respecto a $[a,b]$, o los puntos de $[a,b]$ que no pertenecen a *E*.

Con estas definiciones se puede demostrar una serie de resulta-

[5] Annali di Mat., (3), 7, 1902, 231-259.

dos auxiliares, incluido el hecho de que $m_i(E) \leqslant m_e(E)$. Al conjunto E se le llama por definición medible si se verifica que $m_i(E) = m_e(E)$, y a este valor común se le llama su medida $m(E)$. Lebesgue demuestra que una unión numerable de conjuntos medibles disjuntos dos a dos es medible y que su medida es la suma de las medidas de los conjuntos componentes. También se verifica que todos los conjuntos medibles en el sentido de Jordan lo son en el sentido de Lebesgue y la medida es la misma. El concepto de medida de Lebesgue difiere del de Borel por la adjunción de un conjunto de medida nula en el sentido de Borel. Lebesgue llamó la atención también sobre la existencia de conjuntos no medibles.

El siguiente concepto importante introducido por Lebesgue fue el de función medible. Sea E un conjunto medible acotado del eje Ox; la función $f(x)$, definida en todos los puntos E, se llamará medible en E si el subconjunto formado por los puntos de E tales que $f(x) > A$ es medible, para toda constante A.

Por último, llegamos a la definición de Lebesque de la integral. Sea $f(x)$ una función acotada y medible definida sobre el conjunto medible E contenido en $[a,b]$. Sean A y B los extremos inferior y superior de $f(x)$ sobre E. Dividamos el intervalo $[A,B]$ del eje Oy en n intervalos parciales

$$[A, l_1], [l_1, l_2], \ldots, [l_{n-1}, B],$$

donde $A = l_0$ y $B = l_n$. Sea e_r el conjunto de los puntos de E para los que $l_{r-1} \leqslant f(x) \leqslant l_r$ para $r = 1, 2, 3, \ldots, n$. Entonces $e_1, e_2, \ldots, e_n$ son conjuntos medibles. Construyamos las sumas S y s definidas por

$$S = \sum_1^n l_r m(e_r), \qquad s = \sum_1^n l_{r-1} m(e_r).$$

Estas sumas S y s tienen un extremo inferior J y un extremo superior I respectivamente. Lebesgue demuestra entonces que para toda función acotada medible, $I = J$, y a este valor común se le llama la integral de Lebesgue de $f(x)$ sobre E, y se representa por

$$I = \int_E f(x)\, dx.$$

Si E es todo el intervalo $[a,b]$, entonces se usa la notación $\int_a^b f(x)\,dx$ usual, pero entendiendo que la integral es ahora en el sentido de Lebesgue. Si $f(x)$ es integrable en el sentido de Lebesgue y el valor de la integral es finito, entonces se dice que $f(x)$ es sumable, término introducido por Lebesgue mismo. Toda función $f(x)$ que sea integrable en el sentido de Riemann sobre $[a,b]$ lo es en el sentido de Lebesgue, pero no necesariamente al revés. Si $f(x)$ es integrable en ambos sentidos entonces los valores de las dos integrales coinciden.

La generalidad de la integral de Lebesgue deriva del hecho de que una función integrable para Lebesgue no necesita ser continua casi por doquier (es decir, excepto en un conjunto de medida nula). Así por ejemplo, la función de Dirichlet, que vale 1 para los valores racionales de x y 0 para los valores irracionales del intervalo $[a,b]$ es totalmente discontinua y, aunque no sea en integrable en el sentido de Riemann sí es integrable para Lebesque, y en este caso $\int_a^b(x)\,dx = 0$.

Esta noción de integral de Lebesgue puede extenderse a funciones más generales, por ejemplo a funciones no acotadas. Si $f(x)$ es integrable en el sentido de Lebesgue pero no acotada en el intervalo de integración, la integral converge absolutamente. Las funciones no acotadas pueden ser integrables en el sentido de Lebesgue pero no en el de Riemann y recíprocamente.

Para fines prácticos la integral de Riemann suele ser suficiente. De hecho, Lebesgue mismo demostró (*Leçons sur l'integration et la recherche des fonctions primitives*, 1904) que una función acotada es integrable para Riemann si y sólo si sus puntos de discontinuidad forman un conjunto de medida nula. Pero para la investigación teórica la integral de Lebesgue introduce importantes simplificaciones. Los nuevos teoremas se apoyan en la aditividad numerable de la medida de Lebesgue, en contraste con la aditividad finita del contenido de Jordan.

Para ilustrar la simplicidad de los teoremas utilizando la integral de Lebesgue, veamos un resultado demostrado por Lebesgue mismo en su tesis. Supongamos que $u_1(x)$, $u_2(x)$, ... son funciones sumables sobre un conjunto medible E y que $\Sigma u_n(x)$ converge a $f(x)$; entonces $f(x)$ es medible. Si además $s_n(x) = \Sigma_1^n u_i(x)$ está acotada uniformemente (es decir, si $|s_n(x)| \leq B$ para todo x en E y todo n), entonces es un teorema que $f(x)$ es integrable en el sentido de Lebesgue sobre $[a,b]$ y

$$\int_a^b f(x)\, dx = \lim_{n\to\infty} \int_a^b s_n(x)\, dx.$$

Si hubiéramos estado trabajando con la integral de Riemann, habríamos necesitado la hipótesis adicional de que la suma de la serie fuera integrable; este caso, para la integral de Riemann, es un teorema debido a Cesare Arzelà (1847-1912).[6] Lebesgue hizo de este teorema la piedra angular de la exposición de su teoría y de sus *Leçons sur l'integration.*

La integral de Lebesgue resulta especialmente útil en la teoría de series de Fourier, a la que Lebesgue mismo hizo importantes contribuciones.[7] Según Riemann, los coeficientes de Fourier a_n y b_n de una función acotada e integrable tienden a cero cuando n tiende a infinito. La generalización debida a Lebesgue afirma que

$$\lim_{n\to\infty} \int_a^b f(x) \begin{Bmatrix} \text{sen } nx \\ \cos nx \end{Bmatrix} dx = 0,$$

donde ahora $f(x)$ es cualquier función, acotada o no, que sea integrable en el sentido de Lebesgue. Este resultado se suele denominar hoy lema de Riemann-Lebesgue.

En el mismo artículo de 1903 demuestra Lebesgue que si f es una función acotada representada por una serie trigonométrica, es decir,

$$f(x) = \frac{a_0}{2} + \sum_1^\infty (a_n \cos nx + b_n \text{ sen } nx),$$

entonces los a_n y los b_n son los coeficientes de Fourier. En 1905[8] dio Lebesgue una nueva condición suficiente para la convergencia de la serie de Fourier a la función $f(x)$, que incluía todas las condiciones conocidas previamente.

Lebesgue demostró también (en sus *Leçons sur les séries trigonométriques,* 1906, p. 102) que la posibilidad de integrar término a término una serie de Fourier no depende de la convergencia uniforme de la serie a la función $f(x)$ misma. Lo que ocurre es que

[6] Atti della Accad. dei Lincei, Rendiconti, (4), 1, 1885, 321-326, 532-537, 566-569.
[7] E. g., Ann. de l'Ecole Norm. Sup., (3), 20, 1903, 453-485.
[8] Math. Ann., 61, 1905, 251-280.

$$\int_{-\pi}^{x} f(x)dx = a_0(x+\pi) + \sum_{1}^{\infty} \frac{1}{n} (a_n \text{ sen } nx + b_n (\cos n\pi - \cos nx)),$$

donde x es un punto cualquiera del intervalo $(-\pi,\pi]$, para cualquier $f(x)$ integrable en el sentido de Lebesgue, converja o no la serie original para $f(x)$. Y la nueva serie converge uniformemente al miembro de la izquierda de la ecuación sobre el intervalo $[-\pi,\pi]$.

Además, el teorema de Parseval, en el sentido de que

$$\frac{1}{\pi} \int_{-\pi}^{\pi} [f(x)]^2 dx = 2a_0^2 + \sum_{1}^{\infty} (a_n^2 + b_n^2)$$

se verifica para cualquier $f(x)$ cuyo cuadrado sea integrable sobre $[-\pi,\pi]$ (*Leçons,* 1906, p. 100). Pierre Fatou (1878-1929) demostraría [9] que

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\, dx = 2a_0\alpha_0 + \sum_{1}^{\infty} (a_n\alpha_n + b_n\beta_n),$$

donde a_n, b_n y α_n, β_n son los correspondientes coeficientes de Fourier para $f(x)$ y $g(x)$, cuyos cuadrados son integrables para Lebesgue sobre $[-\pi,\pi]$. A pesar de estos importantes progresos en la teoría de series de Fourier, no se conoce ninguna propiedad de una $f(x)$ integrable en el sentido de Lebesgue en $[-\pi,\pi]$ que sea condición necesaria y suficiente para la convergencia de su serie de Fourier.

Lebesgue dedicó la mayor parte de sus esfuerzos a la conexión entre las nociones de integral y de función primitiva (o integral indefinida). Cuando Riemann introdujo su generalización de la integral, se planteó ya el problema de si la correspondencia entre integral definida y función primitiva, válida para las funciones continuas, se seguía cumpliendo o no en el caso más general. Ahora bien, es posible dar ejemplos de funciones f integrables en el sentido de Riemann y tales que $\int_a^x f(t)\, dt$ no tiene derivada (ni siquiera derivadas laterales, por la derecha o por la izquierda) en algunos puntos. Recíprocamente, Volterra demostró en 1881 [10] que una función $F(x)$ puede tener una derivada acotada en un intervalo I que no sea inte-

---

[9] Acta Math., 30, 1906, 335-400.

[10] Gior di Mat., 19, 1881, 333-372 = *Opere Mat.*, 1, 16-48.

grable en el sentido de Riemann sobre dicho intervalo. Un sutil análisis del problema permitió a Lebesgue demostrar que si $f(x)$ es integrable en su sentido sobre $[a,b]$, entonces $F(x) = \int_a^x f(t)\,dt$ tiene una derivada igual a $f(x)$ casi por doquier, es decir, salvo sobre un conjunto de medida nula *(Leçons sur l'intégration)*. Y recíprocamente, si una función $g(x)$ es diferenciable sobre $[a,b]$ y si su derivada $g' = f$ está acotada, entonces f es integrable en el sentido de Lebesgue y se verifica la fórmula $g(x) - g(a) = \int_a^x f(t)\,dt$. Sin embargo, como advierte el mismo Lebesgue, la situación es mucho más complicada si g' no está acotada. En este caso g' no es necesariamente integrable, y el primer problema es el de caracterizar a las funciones g para las que g' existe casi por doquier y es integrable. Limitándose al caso en que uno de los cuatro números derivados [11] de g sea siempre finito, demuestra Lebesgue que g tiene que ser necesariamente una función de variación acotada (cap. 40, sec. 6). Por último, Lebesgue demostró (en su libro de 1904) el recíproco. Una función g de variación acotada admite una derivada g' casi por doquier, la cual es integrable. Sin embargo, no se tiene necesariamente que

$$g(x) - g(a) = \int_a^x g'(t)\,dt; \qquad (1)$$

sino que la diferencia entre los dos miembros de esta ecuación es una función de variación acotada no constante pero con derivada cero casi por doquier. En cuanto a las funciones de variación acotada g para las que se verifica,[1] tienen la siguiente propiedad: la variación total de g en un conjunto abierto U (es decir, la suma de las variaciones totales de g en cada una de las componentes conexas de U) tiende a cero con la medida de U. A estas funciones las llamó absolutamente continuas Giuseppe Vitali (1875-1932), que además las estudió con detalle.

La obra de Lebesgue también hizo progresar la teoría de las integrales múltiples. Con su definición de integral doble, se amplía el dominio de las funciones para las que la integral doble se puede

[11] Los dos números derivadas por la derecha son los dos límites

$$\limsup_{h\to 0, h>0} \frac{g(x+h) - g(x)}{h}, \qquad \liminf_{h\to 0, h>0} \frac{g(x+h) - g(x)}{h}$$

Los dos números derivadas por la izquierda se definen análogamente.

calcular por integración iterada. Lebesgue ya dio un resultado de este tipo en su tesis de 1902, pero fue mejorado por Guido Fubini (1879-1943):[12] si $f(x,y)$ es sumable sobre el conjunto medible G, entonces

a) $f(x,y)$ como función de x y como función de y es sumable para casi todo y y casi todo x respectivamente;

b) el conjunto de los puntos (x_0,y_0) para los que o bien $f(x,y_0)$ o bien $f(x_0,y)$ no es sumable, tiene medida cero;

$$\text{(c)} \qquad \iint_G f(x,y)\,dG = \int dy\left(\int f(x,y)\,dx\right) = \int dx\left(\int f(x,y)\,dy\right)$$

donde las integrales exteriores están tomadas sobre los conjuntos de puntos y (respectivamente x) para los que $f(x,y)$ como función de x (respectivamente, como función de y) son sumables.

Finalmente, en 1910 [13] llegó Lebesgue a resultados sobre integrales múltiples que generalizaban los de las derivadas para integrales simples. A cada función f integrable en toda región compacta de R^n le asoció la función de conjunto (en oposición a las funciones de variables numéricas) $F(E) = \int_E f(x)\,dx$ (donde x representa aquí n coordenadas) definida para cada dominio de integración E de R^n. Este concepto generaliza la integral indefinida. Lebesgue observó que la función G posee las dos propiedades siguientes:

1) Es completamente aditiva; es decir, $F(\Sigma_1 E_n) = \Sigma_1 F(E_n)$ donde los E_n son conjuntos medibles disjuntos dos a dos.

2) Es absolutamente continua en el sentido de que $F(E)$ tiende a cero con la medida de E.

La parte esencial de este artículo de Lebesgue consistía en demostrar el recíproco de esta proposición, es decir, definir una derivada de $F(E)$ en un punto P del espacio n-dimensional. Lebesgue llegó al siguiente teorema; si $F(E)$ es absolutamente continua y aditiva, entonces tiene una derivada finita casi por doquier, y F es la integral indefinida de la función sumable que es igual a la derivada de F donde ésta exista y sea finita, y arbitraria en los puntos restantes.

La herramienta principal de la demostración es un teorema de recubrimientos debido a Vitali,[14] que sigue siendo fundamental en

[12] Atti della Accad. dei Lincei, Rendiconti, (5), 16, 1907, 608-614.

[13] Ann. de l'Ecole Norm. Sup., (3), 27, 1910, 361-450.

[14] Atti Accad., Torino, 43, 1908, 229-246.

este área de la teoría de integración. Pero Lebesgue no se paró aquí, sino que indicó la posibilidad de generalizar la noción de función de variación acotada considerando funciones $F(E)$, donde E es un conjunto medible, completamente aditivas y tales que $\Sigma_n |F(E_n)|$ permanezca acotada para toda partición numerable de E en subconjuntos medibles E_n. Se podrían citar muchos otros teoremas del cálculo integral que han sido generalizados utilizando el concepto de integral de Lebesgue.

La obra de Lebesgue, una de las grandes contribuciones a la matemática de este siglo, fue aceptada pero, como de costumbre, no sin cierta resistencia. Ya hemos hablado (cap. 40, sec. 7) de las objeciones de Hermite a las funciones sin derivada. El mismo Hermite trató de evitar que Lebesgue publicara una «Nota sobre las superficies no regladas aplicables sobre el plano»[15], en la que Lebesgue estudiaba ciertas superficies no diferenciables. Muchos años más tarde contaba Lebesgue en su *Notice* (p. 14, ver bibliografía del cap.),

> Darboux había dedicado su *Mémoire* de 1875 a la integración y a las funciones sin derivadas, por tanto no experimentaba el mismo horror que Hermite. Sin embargo, dudo de que nunca llegara a perdonarme del todo mi «Nota sobre las superficies aplicables». Seguramente pensaba que los que se enfrascan en estos estudios pierden el tiempo en vez de dedicarse a una investigación útil.

y dice también Lebesgue *(Notice*, p. 13),

> Para muchos matemáticos me convertí en el hombre de las funciones sin derivadas, a pesar de que nunca en la vida me dediqué por completo al estudio o consideración de tales funciones. Y como el temor y el horror que mostraba Hermite lo compartía casi todo el mundo, siempre que intentaba tomar parte en una discusión matemática, aparecía algún analista que decía, «Esto no le interesará a usted; estamos discutiendo de funciones con derivadas». O algún geómetra diciéndolo en su lenguaje: «Estamos discutiendo de superficies que tienen planos tangentes.»

5. Generalizaciones

Ya hemos indicado las ventajas de la integración de Lebesgue tanto para generalizar viejos resultados como para formular nuevos

---

[15] *Comp. Rend.*, 128, 1899, 1502-05.

y elegantes teoremas sobre series. En los capítulos siguientes nos encontraremos con más aplicaciones de las ideas de Lebesgue. Los desarrollos más inmediatos dentro de la teoría de funciones consistieron en diversas extensiones de la idea de integral. De ellas sólo mencionaremos una debida a Johann Radon (1887-1956), que incluye tanto la integral de Stieltjes como la de Lebesgue y que, de hecho, se conoce con el nombre de integral de Lebesgue-Stieltjes.[16] Las generalizaciones consisten no sólo en conceptos de integral distintos y más amplios sobre conjuntos de puntos del espacio euclídeo n-dimensional, sino sobre dominios de espacios más generales, tales como especios de funciones, por ejemplo. Las aplicaciones de estos conceptos más generales pueden encontrarse hoy en la teoría de probabilidades, en teoría espectral, en teoría ergódica y en el análisis armónico (o análisis de Fourier generalizado).

Bibliografía

Borel, Emile: *Notice sur les travaux scientifiques de M. Emile Borel,* Gauthier-Villars, 2.ª ed., 1921.

Bourbaki, Nicolas: *Elementos de historia de las matemáticas,* Madrid, Alianza, 1976.

Collingwood, E. F.: «Emile Borel», Jour. London Math. Soc., 34, 1959, 488-512.

Fréchet, M.: «La vie et l'œuvre d'Emile Borel», L'Enseignement Mathématique, (2), 11, 1965, 1-94.

Hawkins, T. W.; Jr.: *Lebesgue's theory of integration: its origins and development,* University of Wisconsin Press, 1970, caps. 4-6.

Hildebrandt, T. H.: «On integrals related to and extensions of the Lebesgue integral», Amer. Math. Soc. Bull., 24, 1918, 113-177.

Jordan, Camille: *Œuvres,* 4 volúmenes, Paris, Gauthier-Villars, 1961-1964.

Lebesgue, Henri: *Measure and the integral,* Holden-Day, 1966, pp. 176-194. Traducción del francés *Le mesure des grandeurs.*

— *Notice sur les travaux scientifiques de M. Henri Lebesgue,* Edouard Privat, 1922.

— *Leçons sur l'intégration et la recherche des foctions primitives,* Gauthier-Villars, 1904; 2.ª ed., 1928.

McShane, E. J.: «Integrals devised for special purposes»; Amer. Math. Soc. Bull., 69, 1963, 597-627.

---

[16] *Sitzungsber. der Akad. der Wiss. Wien,* 122, Abt, IIa, 1913, 1295-1438.

Pesin, Ivan M.: *Classical and modern integration theories,* Academic Press, 1970.

Plancherel, Michel: «Le développement de la théorie des séries trigonométriques dans le dernier quart de siècle», L'Enseignement Mathématique, 24, 1924-1925, 19-58.

Riesz, F.: «L'evolution de la notion d'intégrale depuis Lebesgue», Annales de l'Institut Fourier, 1, 1949, 29-42.

Capítulo 45
ECUACIONES INTEGRALES

> La naturaleza no se ve desconcertada por las dificultades del análisis.
>
> AUGUSTIN FRESNEL

1. Introducción

Una ecuación integral no es más que una ecuación en la que aparece una función incógnita bajo un signo integral, y el problema de resolver dicha ecuación consiste en determinar esa función. Como vamos a ver enseguida, algunos problemas de la física matemática conducen directamente a ecuaciones integrales, mientras que otros problemas que conducen en primer lugar a ecuaciones diferenciales ordinarias o en derivadas parciales, pueden manejarse a veces con más comodidad convirtiéndolas previamente en ecuaciones integrales. Al principio, a la resolución de ecuaciones integrales se le llamó el problema de la inversión de integrales, y el término ecuación integral lo introdujo Du Bois-Reymond.[1]

Al igual que en otras ramas de la matemática, aparecieron problemas aislados en los que se presentaban ecuaciones integrales mucho antes de que el tema adquiriese su *status* y su metodología propios. Así, por ejemplo, Laplace estudió en 1782 [2] la ecuación integral en $g(t)$ dada por

---

[1] Jour. für Math., 103, 1888, 228.

[2] Mém. de l'Acad. des Sci. Paris, 1782, 1-88, pub. 1785, y 1783, 423-467, pub. 1786 = *Œuvres*, 10, 209-291, p. 236 en particular.

$$f(x) = \int_{-\infty}^{\infty} e^{-xt} g(t)\, dt. \qquad (1)$$

Tal como aparece escrita, la ecuación (1) recibe hoy el nombre de transformación de Laplace de $g(t)$. Poisson [3] descubrió la expresión de $g(t)$, a saber,

$$g(t) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{xt} f(x)\, dx$$

para a suficientemente grande. Otro de los resultados notables que en realidad pertenecen a la historia de las ecuaciones integrales proviene del famoso artículo de Fourier de 1811 sobre la propagación del calor (cap. 28, párr. 3). Allí nos encontramos con la ecuación

$$f(x) = \int_{0}^{\infty} \cos (xt) u(t)\, dt$$

y la fórmula de inversión

$$u(t) = \frac{2}{\pi} \int_{0}^{\infty} \cos (xt) f(x)\, dx.$$

El primer caso de uso directo y consciente de una ecuación integral y su resolución se remontan a Abel. En dos de sus primeros artículos publicados, el primero en 1823 [4] en una oscura revista y el segundo en el *Journal für Mathematik,*[5] considera Abel el siguiente problema de mecánica: una partícula material cae desde P deslizándose por una curva lisa (fig. 45.1) hasta el punto O; la curva está situada en un plano vertical. La velocidad adquirida en O es independiente de la forma de la curva, pero el tiempo empleado en deslizar desde P hasta O no lo es. Si llamamos (ξ,η) a las coordenadas de un punto genérico Q entre P y Q, y s es el arco PQ, entonces la velocidad de la partícula en Q vendrá dada por

---

[3] Jour. de l'Ecole Poly., 1823, 1-144, 249-403.
[4] Magazin for Naturwidenskaberne, 1, 1823 = *Œuvres*, 1, 11-27.
[5] Jour. für Math., 1, 1826, 153-157 = *Œuvres*, 1, 97-101.

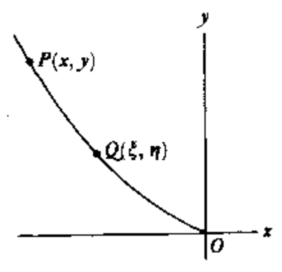


FIGURA 45.1

$$\frac{ds}{dt} = -\sqrt{2g(x-\xi)}$$

donde g es la constante gravitatoria. Por tanto

$$t = \frac{-1}{\sqrt{2g}} \int_P^Q \frac{ds}{\sqrt{x-\xi}}.$$

Ahora bien, s puede expresarse en términos de ξ. Supongamos que s es $v(\xi)$. Entonces el tiempo total de descenso T desde P hasta O vendrá dado por

$$T(x) = \frac{1}{\sqrt{2g}} \int_0^x \frac{v'(\xi)\, d\xi}{\sqrt{x-\xi}}.$$

Claramente, el tiempo T depende de x para cualquier curva. El problema que se propone Abel es, dado T como función de x, hallar $v(\xi)$. Si introducimos

$$f(x) = \sqrt{2g}\, T(x)$$

el problema se convierte en el de determinar v en la ecuación

$$f(x) = \int_0^x \frac{v'(\xi)}{\sqrt{x-\xi}}\, d\xi.$$

Abel obtiene la solución

$$v(\xi) = \int_0^\xi \frac{f(x)\, dx}{(\xi - x)^{1/2}}.$$

Los métodos que utiliza (da dos) son particulares y no nos interesan ahora.

En realidad, de lo que trata Abel es de resolver el problema más general

$$f(x) = \int_a^x \frac{u(\xi)\, d\xi}{(x-\xi)^\lambda}, \qquad 0 < \lambda < 1 \tag{2}$$

y obtiene

$$u(z) = \frac{\operatorname{sen} \lambda\pi}{\pi} \frac{d}{dz} \int_a^z \frac{f(x)\, dx}{(z-x)^{1-\lambda}}.$$

Liouville, que trabajaba independientemente de Abel, resolvió algunas ecuaciones integrales concretas desde 1832 [6] en adelante. Un paso más importante dado por Liouville [7] fue el de mostrar cómo se puede obtener la solución de ciertas ecuaciones diferenciales resolviendo ecuaciones integrales. La ecuación diferencial que había que resolver era la

$$y'' + [\rho^2 - \sigma(x)]y = 0 \tag{3}$$

sobre el intervalo $a \leqslant x \leqslant b$, con ρ como parámetro. Sea $u(x)$ la solución particular que satisface las condiciones iniciales

$$u(a) = 1, \qquad u'(a) = 0. \tag{4}$$

Esta función también será solución de la ecuación no homogénea

$$y'' + \rho^2 y = \sigma(x)u(x).$$

Entonces, por un resultado básico de ecuaciones diferenciales

$$u(x) = \cos \rho(x-a) + \frac{1}{\rho}\int_a^x \sigma(\xi) \operatorname{sen} \rho(x-\xi) u(\xi)\, d\xi. \tag{5}$$

Así pues, si podemos resolver esta ecuación integral habremos

[6] Jour. de l'Ecole Poly., 13, 1832, 1-69.
[7] Jour. de Math., 2, 1837, 16-35.

obtenido la solución de la ecuación diferencial (3) que satisface las condiciones iniciales (4).

Liouville obtuvo la solución por un método de sustituciones sucesivas más tarde atribuido a Carl G. Neumann, cuya obra *Untersuchungen über das logarithmische und Newton'sche Potential* (1877) apareció treinta años después. No explicaremos el método de Liouville porque es prácticamente idéntico al de Volterra, del que hablaremos más adelante.

Las ecuaciones integrales estudiadas por Abel y por Liouville son de dos tipos básicos; la de Abel es de la forma

$$f(x) = \int_a^x K(x, \xi)u(\xi)\, d\xi, \tag{6}$$

y la de Liouville de la forma

$$u(x) = f(x) + \int_a^x K(x, \xi)u(\xi)\, d\xi. \tag{7}$$

En ambas $f(x)$ y $K(x,\xi)$ son funciones conocidas y $u(\xi)$ es la función que hay que determinar. La terminología utilizada hoy, que fue introducida por Hilbert, se refiere a estas ecuaciones llamándolas de primero y segundo tipo respectivamente, y $K(x,\xi)$ se denomina el núcleo de la ecuación. Como hemos dicho ya, también reciben el nombre de ecuaciones de Volterra, mientras que cuando el límite superior de integración es un número fijo b, se las llama ecuaciones de Fredholm. En realidad, las ecuaciones de Volterra son a su vez casos particulares, respectivamente, de las de Fredholm, porque siempre se puede tomar $K(x,\xi) = 0$ para $\xi > x$ y considerar las ecuaciones de Volterra como ecuaciones de Fredholm. El caso especial de ecuación del segundo tipo en que $f(x) \equiv 0$ recibe el nombre de ecuación homogénea.

A mediados del siglo XIX el principal interés en las ecuaciones integrales se centraba en torno a la resolución de los problemas de condiciones de contorno asociados a la ecuación del potencial

$$\Delta u = u_{xx} + u_{yy} = 0. \tag{8}$$

La ecuación se verifica en una región plana dada, limitada por una curva C. Si el valor en la frontera de u es alguna función $f(s)$

dada en términos de la longitud de arco s a lo largo de C, entonces una solución de este problema del potencial viene dada por

$$u(x, y) = \frac{1}{2\pi} \int_C \varrho(s) \log \frac{1}{r(s; x, y)} \, ds,$$

donde $r(s;x,y)$ es la distancia de un punto s de C a un punto cualquier (x,y) del interior o de la frontera, y $\varrho(s)$ es una función desconocida que verifica, para $s = (x,y)$ sobre C

$$f(s) = \frac{1}{2\pi} \int_C \varrho(t) \log \frac{1}{r(t; x, y)} \, dt. \tag{9}$$

Esta es una ecuación integral de primer tipo para $\varrho(t)$. De manera alternativa, si uno toma como solución de (8) con la misma condición de frontera

$$v(x, y) = \frac{1}{2\pi} \int_C \phi(s) \frac{\partial}{\partial n} \left(\log \frac{1}{r(s; x, y)} \right) ds,$$

donde $\partial/\partial n$ representa la derivada normal en la frontera, entonces $\phi(s)$ tiene que satisfacer la ecuación integral

$$f(s) = \frac{1}{2} \phi(s) + \frac{1}{2\pi} \int_C \phi(t) \frac{\partial}{\partial n} \left(\log \frac{1}{r(t; x, y)} \right) dt, \tag{10}$$

que es una ecuación integral del segundo tipo. Neumann resolvió estas ecuaciones para regiones convexas en sus *Untersuchungen* y en publicaciones posteriores.

Otro problema de ecuaciones en derivadas parciales fue atacado por medio de ecuaciones integrales; la ecuación

$$\Delta u + \lambda u = f(x, y) \tag{11}$$

aparece en el estudio de movimientos de ondas cuando la dependencia temporal de la correspondiente ecuación hiperbólica

$$\Delta u - \frac{1}{c^2} u_{tt} = f(x, y),$$

que suele tomarse $e^{-i\omega t}$, se elimina. Se sabía (cap. 28, sec. 8) que el

caso homogéneo de la ecuación (11) sometido a las condiciones de contorno tiene soluciones no triviales sólo para un conjunto discreto de valores de λ, llamados autovalores o valores característicos. En 1894 [8] estudió Poincaré el caso no homogéneo de (11) con λ complejo, y logró construir una función meromorfa en λ que representaba la solución única de (11) para cualquier λ que no fuese un autovalor, y cuyos residuos producían autofunciones para el caso homogéneo, es decir, cuando $f = 0$.

Sobre la base de estos resultados, consideró Poincaré en 1896 [9] la ecuación

$$u(x) + \lambda \int_a^b K(x, y)u(y)\, dy = f(x),$$

que obtuvo de (11), afirmando que la solución sería una función meromorfa en λ. Este resultado lo demostraría Fredholm en un artículo que vamos a comentar en breve.

La transformación de ecuaciones diferenciales en ecuaciones integrales, que hemos ilustrado con los anteriores ejemplos, se convirtió en una técnica importante para resolver problemas de contorno en ecuaciones diferenciales ordinarias y en derivadas parciales, y constituyó el más fuerte impulso para el estudio de las ecuaciones integrales en sí mismas.

2. Los comienzos de una teoría general

Vito Volterra (1860-1940), sucesor de Beltrami como profesor de física matemática en Roma, fue el primero de los fundadores de una teoría general para las ecuaciones integrales. Escribió diversos artículos sobre el tema desde 1884, de los cuales los más importantes datan de 1896 y 1897.[10] Volterra ideó un método para resolver ecuaciones integrales de segundo tipo,

---

[8] Rendiconti del Circolo Matematico di Palermo, 8, 1894, 57-155 = *Œuvres*, 9, 123-196.

[9] Acta Math., 20, 1896-1897, 59-142 = *Œuvres*, 9, 202-272. Véase también la referencia de Hellinger y Toeplitz en la bibliografía, p. 1354.

[10] Atti della Acad. dei Lincei, Rendiconti, (5), 1896, 177-185, 289-300; Atti Accad. Torino, 31, 1896, 311-323, 400-408, 557-567; Annali di Mat., (2), 25, 1897, 139-178; todos están en sus *Opere matematiche*, 2, 216-313.

$$f(s) = \phi(s) + \int_a^b K(s, t)\phi(t)\, dt \tag{12}$$

donde $\phi(s)$ es la incógnita y $K(s,t) = 0$ para $t > s$. Volterra escribe esta ecuación de la forma

$$f(s) = \phi(s) + \int_a^s K(s, t)\phi(t)\, dt.$$

y su método de resolución consiste en definir

$$\begin{aligned} f_1(s) &= -\int_a^b K(s, t)f(t)\, dt \\ &\cdots\cdots\cdots\cdots \\ f_n(s) &= -\int_a^b K(s, t)f_{n-1}(t)\, dt \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{13}$$

y tomar $\phi(s)$ como

$$\phi(s) = f(s) + \sum_{p=1}^{\infty} f_p(s). \tag{14}$$

Para su núcleo $K(s,t)$ pudo demostrar Volterra la convergencia de la serie (14), y si uno sustituye (14) en (12) puede ver que es una solución. Esta sustitución da

$$\phi(s) = f(s) + \int_a^b \overline{K}(s, t)f(t)\, dt, \tag{15}$$

que puede escribirse en la forma

$$\phi(s) = f(s) - \int_a^b K(s, t)f(t)\, dt + \int_a^b \int_a^b K(s, r)K(r, t)f(t)\, dr\, dt + \ldots,$$

donde ahora el núcleo $\bar{K}$ (llamado más tarde núcleo resolvente o simplemente resolvente por Hilbert) es

$$\overline{K}(s, t) = -K(s, t) + \int_a^b K(s, r)K(r, t)\, dr - \int_a^b \int_a^b K(s, r)K(r, w)K(w, t)\, dr\, dw + \ldots$$

La ecuación (15) es la representación obtenida anteriormente por Liouville para una ecuación integral particular, y atribuida a Neumann. Volterra resolvió también ecuaciones integrales de primer tipo $f(s) = \int_a^s K(x,s)\, \phi(x)\, dx$ reduciéndolas a ecuaciones del segundo tipo.

En 1896 observó Volterra que una ecuación integral del primer tipo venía a ser una forma límite de un sistema de n ecuaciones lineales con n incógnitas, cuando n tiende a infinito. Erik Ivar Fredholm (1866-1927), profesor de matemáticas en Estocolmo, interesado en la resolución del problema de Dirichlet, recogió esta idea en 1900 [11] y la utilizó para resolver ecuaciones integrales del segundo tipo, es decir, ecuaciones de la forma (12), sin la restricción sobre $K(s,t)$ sin embargo.

Escribiremos la ecuación que estudió Fredholm en la forma

$$u(x) = f(x) + \lambda \int_a^b K(x, \xi)u(\xi)\, d\xi, \tag{16}$$

aunque el parámetro λ no aparecía explícitamente en su obra; sin embargo, es más fácil de entender lo que él hizo, a la luz de trabajos posteriores, si lo hacemos explícito. Para ajustarnos fielmente a las fórmulas de Fredholm bastaría hacer $\lambda = 1$ o bien suponerlo implícitamente incluido en K.

Fredholm divide el intervalo en x $[a,b]$ en n partes iguales por medio de los puntos

$$a,\ x_1 = a + \delta,\ x_2 = a + 2\delta,\ \ldots,\ x_n = a + n\delta = b.$$

y a continuación reemplaza la integral que figura en (16) por la suma

$$u_n(x) = f(x) + \sum_{j=1}^{n} \lambda K(x, x_j)u_n(x_j)\delta. \tag{17}$$

[11] Acta Math., 27, 1903, 365-390.

Ahora bien, la ecuación (17) se supone que se verifica para todos los valores de x del intervalo $[a,b]$. Por tanto se debe verificar para $x = x_1, x_2, ..., x_n$. Esto conduce al sistema de n ecuaciones

$$-\sum_{j=1}^{n} \lambda K(x_i, x_j) u_n(x_j)\delta + u_n(x_i) = f(x_i), \qquad i = 1, 2, ..., n. \quad (18)$$

Este es un sistema lineal no homogéneo de n ecuaciones para determinar las n incógnitas $u_n(x_1)$, $u_n(x_2)$, ... $u_n(x_n)$.

En la teoría de sistemas de ecuaciones lineales se conocía ya el siguiente resultado. Si tenemos una matriz

$$S_n = \left\| \begin{matrix} 1 + a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 1 + a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 + a_{nn} \end{matrix} \right\|$$

entonces el determinante $D(n)$ de S_n admite el siguiente desarrollo

$$D(n) = 1 + \frac{1}{1!}\sum_{r_1} a_{r_1 r_1} + \frac{1}{2!}\sum_{r_1, r_2} \begin{vmatrix} a_{r_1 r_1} & a_{r_1 r_2} \\ a_{r_2 r_1} & a_{r_2 r_2} \end{vmatrix} + \dots$$

$$+ \frac{1}{n!}\sum_{r_1, \dots, r_n} \begin{vmatrix} a_{r_1 r_1} & a_{r_1 r_2} & \dots & a_{r_1 r_n} \\ a_{r_2 r_1} & a_{r_2 r_2} & \dots & a_{r_2 r_n} \\ \dots & \dots & \dots & \dots \\ a_{r_n r_1} & a_{r_n r_2} & \dots & a_{r_n r_n} \end{vmatrix}$$

donde $r_1, r_2, ..., r_n$ recorren independientemente todos los valores de 1 a n. Desarrollando el determinante de los coeficientes en (18) y haciendo tender después n a infinito, obtuvo Fredholm el determinante

$$D(\lambda) = 1 - \lambda \int_a^b K(\xi_1, \xi_1)\, d\xi_1$$
$$+ \frac{\lambda^2}{2!}\int_a^b \int_a^b \begin{vmatrix} K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} d\xi_1\, d\xi_2 - \dots \quad (19)$$

[12] Puede encontrarse una excelente exposición en Gerhard Kowaleski, *Integralgleichungen*, Walter de Gruyter, 1930, 101-134.

al que llamó determinante de la ecuación (16) o del núcleo K. Análogamente, considerando los adjuntos de los elementos de la fila μ y de la columna ν del determinante de los coeficientes en (18) y haciendo tender n al infinito, obtuvo Fredholm la función

$$\begin{aligned} D(x, y, \lambda) = \lambda K(x, y) - \lambda^2 \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) \\ K(\xi_1, y) & K(\xi_1, \xi_1) \end{vmatrix} d\xi_1 \\ + \frac{\lambda^3}{2} \int_a^b \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) & K(x, \xi_2) \\ K(\xi_1, y) & K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, y) & K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} d\xi_1\, d\xi_2 - \ldots \end{aligned} \tag{20}$$

Fredholm llamó a esta función $D(x,y,\lambda)$ el primer menor del núcleo K porque juega un papel análogo al de los primeros menores en el caso de n ecuaciones lineales con n incógnitas. También llamó a los ceros de la función analítica entera $D(\lambda)$ las raíces de $K(x,y)$. Aplicando la regla de Cramer al sistema de ecuaciones lineales (18) y haciendo tender n a infinito, fue como dedujo Fredholm la forma de la solución de (16). A continuación demostró que era correcta sustituyéndola directamente, y pudo al fin enunciar el siguiente resultado: si λ no es una de las raíces de K, es decir, si $D(\lambda)$ es distinto de cero, entonces (16) tiene una y sólo una solución (continua), a saber,

$$u(x, \lambda) = f(x) + \int_a^b \frac{D(x, y, \lambda)}{D(\lambda)} f(y)\, dy. \tag{21}$$

Por otra parte, si λ es una raíz de $K(x,y)$, entonces (16) o bien no tiene ninguna solución continua o tiene infinitas.

Fredholm obtuvo además resultados acerca de la relación entre la ecuación homogénea

$$u(x) = \lambda \int_a^b K(x, \xi) u(\xi)\, d\xi \tag{22}$$

y la no homogénea (16). Es casi evidente a partir de (21) que cuando λ no sea raíz de K la única solución continua de (22) es $u \equiv 0$; por tanto, se dedicó a estudiar el caso en que λ es raíz de K. Sea $\lambda = \lambda_1$ una tal raíz. Entonces (22) tiene las infinitas soluciones

$$c_1u_1(x) + c_2u_2(x) + ... + c_nu_n(x),$$

donde las c_i son constantes arbitrarias; las u_1, u_2, ..., u_n, llamadas soluciones principales, son linealmente independientes, y n depende de λ_1. El número n recibe el nombre de índice de λ_1 (que no es la multiplicidad de λ_1 como cero de $D(\lambda)$). Fredholm pudo calcular el índice de cualquier raíz λ_i y demostrar que el índice nunca puede exceder de la multiplicidad (que siempre es finita). Las raíces de $D(\lambda) = 0$ reciben el nombre de valores característicos de $K(x,y)$ y al conjunto de las raíces se le llama su espectro. Las soluciones de (22) correspondientes a los valores característicos reciben el nombre de autofunciones o funciones características.

Ahora estaba ya Fredholm en condiciones de establecer lo que ha sido llamado desde entonces el teorema de la alternativa de Fredholm. En el caso en que λ sea un valor característico de K, no solamente la ecuación integral (22) tiene n soluciones independientes, sino que la ecuación asociada o adjunta, que tiene el núcleo transpuesto, es decir,

$$u(x) = \lambda \int_a^b K(\xi, x)u(\xi)\, d\xi,$$

también tiene n soluciones $\psi_1(x)$, $\psi_2(x)$, ..., $\psi_n(x)$ para el mismo valor característico, y entonces la ecuación no homogénea (16) es soluble si y sólo si

$$\int_a^b f(x)\psi_i(x)\, dx = 0, \qquad i = 1, 2, ..., n. \tag{23}$$

Estos últimos resultados siguen un paralelismo muy estrecho con la teoría de sistemas de ecuaciones lineales, homogéneos y no homogéneos.

3. La obra de Hilbert

El interés de Hilbert por las ecuaciones integrales se despertó oyendo una conferencia dada por Erik Holmgren (n. 1872) en 1901 sobre la obra de Fredholm en este campo, que ha había sido publicada en Suecia. David Hilbert (1862-1943), uno de los más impor-

tantes matemáticos de este siglo, que había realizado ya una obra impresionante sobre números algebraicos, teoría de invariantes y los fundamentos de la geometría, volvía ahora su atención a las ecuaciones integrales. El mismo nos dice que una investigación sobre el tema le mostró que era importante para la teoría de integrales definidas, para el desarrollo de funciones arbitrarias en series (de funciones especiales o de funciones trigonométricas), para la teoría de ecuaciones diferenciales lineales, para la teoría del potencial y para el cálculo de variaciones. Desde 1904 a 1910 escribió una serie de seis artículos en el *Nachrichten von der Köninglichen Gesellschaft der Wissenschaften zu Göttingen*, artículos recogidos posteriormente en su libro *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (1912). En la última parte de esta obra aplica las ecuaciones integrales a problemas de la física matemática.

Fredholm había utilizado la analogía existente entre las ecuaciones integrales y las ecuaciones algebraicas lineales, pero en vez de llevar a cabo el proceso de paso al límite para las infinitas ecuaciones algebraicas resultantes, se limitó a escribir audazmente los determinantes que salían de esa analogía y a comprobar que resolvían la ecuación integral. Lo primero que hizo Hilbert entonces fue llevar a cabo rigurosamente el paso al límite sobre el sistema finito de ecuaciones lineales.

Hilbert parte de la ecuación integral

$$f(s) = \xi(s) - \lambda \int_0^1 K(s, t)\phi(t)\, dt, \tag{24}$$

donde $K(s,t)$ es continua. El parámetro λ aparece explícitamente y juega de hecho un papel importante en la teoría que sigue. Al igual que Fredholm, Hilbert divide el intervalo [0,1] en n partes, de manera que p/n y q/n ($p,q = 1, 2, \ldots, n$) representan puntos del intervalo [0,1]. Sean

$$K_{pq} = K\left(\frac{p}{n}, \frac{q}{n}\right), \qquad f_p = f\left(\frac{p}{n}\right), \qquad \phi_q = \phi\left(\frac{q}{n}\right).$$

Entonces, de (24) obtenemos el sistema de n ecuaciones con n incógnitas $\phi_1, \phi_2, \ldots, \phi_n$, es decir,

$$f_p = \phi_p - \lambda \sum_{q=1}^{n} K_{pq}\phi_q, \qquad p = 1, 2, \ldots, n.$$

Después de hacer un repaso de la teoría general de la resolución de los sistemas finitos de n ecuaciones lineales con n incógnitas, Hilbert vuelve a considerar directamente la ecuación (24). Para el núcleo K de (24) los autovalores están definidos como los ceros de la serie de potencias

$$\delta(\lambda) = 1 + \sum_{n=1}^{\infty} (-1)^n d_n \lambda^n,$$

donde los coeficientes d_n vienen dados por

$$d_n = \frac{1}{n!} \int_0^1 \dots \int_0^1 |\{K(s_i, s_j)\}| \, ds_1 \dots ds_n.$$

Aquí $|\{K(s_i,s_j)\}|$ es el determinante de la matriz $n \times n$ $\{K(s_i,s_j)\}$, $i,j = 1, 2, \dots, n$, y los s_i son valores de t en el intervalo [0,1]. Para explicar el resultado principal obtenido por Hilbert necesitamos definir las siguientes funciones

$$\Delta_p(x, y) = \frac{1}{p!} \int_0^1 \dots \int_0^1 \begin{vmatrix} 0 & x(s_1) & \dots & x(s_p) \\ y(s_1) & K(s_1, s_2) & \dots & K(s_1, s_p) \\ \vdots & \vdots & \vdots & \vdots \\ y(s_p) & K(s_p, s_1) & \dots & K(s_p, s_p) \end{vmatrix} ds_1 \dots ds_p,$$

donde $x(r)$ e $y(r)$ son funciones continuas arbitrarias cuando r recorre el intervalo [0,1], y

$$\Delta(\lambda; x, y) = \sum_{p=1}^{\infty} (-1)^p \Delta_p(x, y) \lambda^{p-1}.$$

Hilbert define a continuación

$$\Delta^*(\lambda; s, t) = \lambda \Delta(\lambda; x, y) - \delta(\lambda)$$

donde ahora $x(r) = K(s,r)$ e $y(r) = K(r,t)$, y demuestra que si $\bar{K}$ se define como

$$\overline{K}(s, t) = \frac{\Delta^*(\lambda; s, t)}{-\delta(\lambda)}$$

para valores de λ tales que $\delta(\lambda) \neq 0$, entonces

$$K(s, t) = \overline{K}(s, t) - \lambda \int_0^1 \overline{K}(s, r)K(r, t)\, dr$$

$$= \overline{K}(s, t) - \lambda \int_0^1 K(s, r)\overline{K}(r, t)\, dr.$$

Por último, si se toma ϕ definida por

$$\phi(r) = f(r) + \lambda \int_0^1 \overline{K}(r, t)f(t)\, dt, \qquad (25)$$

ϕ es una solución de (24). Las demostraciones de las diversas etapas de esta teoría requieren un cierto número de consideraciones de paso al límite sobre expresiones que aparecían en el tratamiento de Hilbert de los sistemas de ecuaciones lineales.

Hasta aquí Hilbert había demostrado que para cualquier núcleo continuo (no necesariamente simétrico) $K(s,t)$ y para cualquier valor de λ tal que $\delta(\lambda) \neq 0$, existe la función resolvente $\bar{K}(s,t)$ con la propiedad de que la función (25) es solución de la ecuación (24).

A continuación Hilbert supone que $K(s,t)$ es simétrico, lo que le permite hacer uso de propiedades de las matrices simétricas en el caso finito, y demuestra que los ceros de $\delta(\lambda)$, es decir, los autovalores del núcleo simétrico son reales. Entonces los ceros de $\delta(\lambda)$ se ordenan en valor absoluto creciente (para valores absolutos iguales el cero positivo se toma primero, y han de tenerse en cuenta también las multiplicidades de las raíces). Las autofunciones de la ecuación (24) vienen definidas ahora por

$$\phi^k(s) = \left(\frac{\lambda_k}{\Delta^*(\lambda_k; s^*, s^*)}\right)^{1/2} \Delta^*(\lambda_k; s, s^*),$$

donde s^* se toma de manera que $\Delta^*(\lambda_k; s^*,s^*) \neq 0$ y λ_k es un autovalor cualquiera de $K(s,t)$.

Las autofunciones asociadas a los distintos autovalores pueden tomarse formando un sistema ortonormal (ortogonal y normaliza-

do),[13] y para cada autovalor λ_k y cada autofunción correspondiente a λ_k

$$\phi^k(s) = \lambda_k \int_0^1 K(s, t)\phi^k(t)\, dt.$$

Con estos resultados Hilbert puede demostrar ya lo que se conoce como teorema del eje principal generalizado para las formas cuadráticas simétricas. Sea

$$\sum_{p=1}^{n} \sum_{q=1}^{n} k_{pq}x_p x_q \tag{26}$$

una forma cuadrática n-dimensional en las n variables $x_1, x_2, ..., x_n$. Podemos escribirla en la forma (Kx, x), donde K es la matriz de los k_{pq}, x representa el vector $(x_1, x_2, ..., x_n)$ y (Kx,x) es el producto escalar de los vectores Kx y x. Supongamos que K tiene los n autovalores distintos $\lambda_1, \lambda_2, ... \lambda_n$; entonces, para cada λ_k fijo, las ecuaciones

$$0 = \phi_p - \lambda_p \sum_{q=1}^{n} k_{pq}\phi_q, \qquad p = 1, 2, ..., n \tag{27}$$

tienen la solución

$$\phi^k = (\phi_1^k, \phi_2^k, ..., \phi_n^k),$$

que es única salvo un factor constante. Se tiene entonces, tal como demostró Hilbert

$$K(x, x) = \sum_{k=1}^{n} \frac{1}{\lambda_k} \frac{(\phi^k, x)^2}{(\phi^k, \phi^k)} \tag{28}$$

donde los paréntesis representan de nuevo el producto escalar de dos vectores.

El teorema generalizado de los ejes principales de Hilbert puede enunciarse de la manera siguiente: sea $K(s,t)$ una función continua

[13] Normalización significa que se modifica $\phi^k(s)$ de tal modo que se tenga $\int_0^1 (\phi^k)^2 ds = 1$.

simétrica de s y t. Sea $\phi^p(s)$ la autofunción normalizada correspondiente al autovalor λ_p de la ecuación integral (24). Entonces se verifica la siguiente relación, donde $x(s)$ e $y(s)$ son funciones continuas arbitrarias,

$$\int_a^b\int_a^b K(s,t)x(s)y(t)dsdt=\sum_{p=1}^{\alpha}\frac{1}{\lambda_p}\left(\int_a^b\phi^p(s)x(s)ds\right)\left(\int_a^b\phi^p(s)y(s)ds\right), \tag{29}$$

donde $\alpha = n$ o ∞, dependiendo del número de autovalores, y en el segundo caso la suma converge uniforme y absolutamente para todo $x(s)$ e $y(s)$ que satisfagan las condiciones

$$\int_a^b x^2(s)\,ds < \infty \quad \text{y} \quad \int_a^b y^2(s)\,ds < \infty.$$

La generalización de (28) a (29) se pone de manifiesto si definimos en primer lugar $\int_a^b u(s)\,v(s)\,ds$ como el producto escalar de las dos funciones $u(s)$ y $v(s)$ y lo representamos por (u,v). Ahora basta reemplazar $y(s)$ en (29) por $x(s)$ y en el miembro de la derecha de (28) la suma por integración.

A continuación demostró Hilbert un famoso resultado que más tarde recibió el nombre de teorema de Hilbert-Schmidt. Si $f(s)$ es tal que para alguna función continua $g(s)$ se verifica que

$$f(s) = \int_a^b K(s,t)g(t)\,dt \tag{30}$$

entonces

$$f(s) = \sum_{p=1} c_p\phi^p, \tag{31}$$

donde las ϕ^p son las autofunciones ortonormales de K y

$$c_p = \int_a^b \phi^p(s)f(s)\,ds. \tag{32}$$

Así pues, una función «arbitraria» $f(s)$ puede expresarse como un desarrollo en serie en las autofunciones de K con coeficientes c_p que son los «coeficientes de Fourier» del desarrollo.

En el trabajo anterior, Hilbert llevó a cabo un proceso de paso al límite que le permitió generalizar resultados sobre sistemas finitos de ecuaciones lineales y formas cuadráticas finitas a integrales y ecuaciones integrales. Sobre esta base decidió que un tratamiento de las formas cuadráticas infinitas, es decir, las formas cuadráticas con infinitas variables, por sí mismas, «vendría a completar de una manera esencial la teoría bien conocida de las formas cuadráticas con un número finito de variables», lo cual le llevó a ocuparse de problemas que pueden ser considerados como puramente algebraicos. Partiendo de una forma bilineal infinita

$$K(x, y) = \sum_{p,q=1}^{\infty} k_{pq} x_p y_q,$$

y procediendo por paso al límite sobre resultados para formas bilineales y cuadráticas en $2n$ y n variables respectivamente, obtiene Hilbert algunos resultados básicos; los detalles del trabajo son complicados y nos limitaremos a formular solamente algunos de estos resultados. En primer lugar obtiene Hilbert una expresión para una forma resolvente $\bar{K}(\lambda; x, x)$ que tiene la característica peculiar de que es la suma de varias expresiones, una para cada valor de λ de un cierto conjunto discreto, y de una integral sobre un conjunto de λ pertenecientes a un dominio continuo. El conjunto discreto de valores de λ corresponde al espectro puntual de K, y el conjunto continuo al espectro continuo o de banda. Esta es la primera aplicación importante de los espectros continuos, que habían sido observados ya en 1896 para ecuaciones en derivadas parciales por Wilhelm Wirtinger (n. 1865).[14]

Para llegar al resultado clave sobre formas cuadráticas, introdujo Hilbert el concepto de forma acotada. La notación (x,x) representa el producto escalar del vector $(x_1, x_2, ..., x_n, ...)$ consigo mismo, y (x,y) el producto análogo de los vectores x e y. Entonces, la forma $K(x,y)$ se llama acotada si $|K(x,y)| \leqslant M$ para todo x e y tales que $(x,x) \leqslant 1$ e $(y,y) \leqslant 1$. La acotación implica la continuidad, que define Hilbert para una función de infinitas variables.

El resultado clave de Hilbert que hemos mencionado es la generalización a formas cuadráticas con infinitas variables del conocido teorema de los ejes principales de la geometría analítica. Hil-

[14] Math. Ann., 48, 1897, 365-389.

bert demuestra que existe una transformación ortogonal T tal que en las nuevas variables x'_p, donde $x' = Tx$, la forma K se reduce a una «suma de cuadrados». Es decir, toda forma cuadrática acotada $K(x,x) = \Sigma^{\infty}_{p,q=1} k_{pq} x_p x_q$ puede transformarse mediante una transformación ortogonal única en la forma

$$K(x, x) = \sum_{i=1}^{\infty} k_i x_i^2 + \int_{(s)} \frac{d\sigma(\mu, \xi)}{\mu}, \tag{33}$$

donde los k_i son los autovalores recíprocos de K. La integral, en cuya explicación no entraremos, está extendida a un dominio continuo de autovalores o espectro continuo.

Para eliminar el espectro continuo introduce Hilbert el concepto de continuidad completa: una función $F(x_1, x_2, ...)$ de infinitas variables se llama completamente continua en el punto $a = (a_1,a_2, ...)$ si

$$\lim_{\substack{\varepsilon_1 \to 0,\\ \varepsilon_2 \to 0,\\ \dots}} F(a_1 + \varepsilon_1, a_2 + \varepsilon_2, ...) = F(a_1, a_2, ...)$$

siempre que ε_1, ε_2, ... recorran cualquier sistema de valores $\varepsilon_1^{(h)}$, $\varepsilon_2^{(h)}$, ... con límites

$$\lim_{h\to\infty} \varepsilon_1^{(h)} = 0, \lim_{h\to\infty} \varepsilon_2^{(h)} = 0, ...$$

Esta es una condición más fuerte que la continuidad introducida previamente por Hilbert.

Para que una forma cuadrática $K(x,x)$ sea completamente continua es suficiente que $\Sigma^{\infty}_{p,q=1} k^2_{pq} < \infty$. Con esta condición adicional pudo demostrar Hilbert que si K es una forma acotada completamente continua entonces, mediante una transformación ortogonal, puede reducirse a la forma

$$K(x, x) = \sum_j k_j x_j^2, \tag{34}$$

donde los k_j son autovalores recíprocos y $(x_1, x_2, ...)$ satisface la condición de que $\Sigma_1^{\infty} x^2$ es finita.

A continuación Hilbert se dedica a aplicar su teoría de formas cuadráticas en infinitas variables a las ecuaciones integrales. Los re-

sultados no son nuevos, la mayoría de ellos, pero se obtienen por métodos más claros y sencillos. Hilbert comienza esta nueva etapa de su trabajo sobre ecuaciones integrales definiendo el importante concepto de sistema ortogonal completo e funciones $\{\phi_p(s)\}$. Se trata de una sucesión de funciones definidas y continuas todas ellas sobre el intervalo $[a,b]$, con las siguientes propiedades:

a) ortogonalidad:

$$\int_a^b \phi_p(s)\phi_q(s) = \delta_{pq}, \qquad p, q = 1, 2, \ldots$$

b) completitud: para todo par de funciones u y v definidas sobre el intervalo $[a,b]$

$$\int_a^b u(s)v(s)\,ds = \sum_{p=1}^{\infty} \int_a^b \phi_p(s)u(s)\,ds \int_a^b \phi_p(s)v(s)\,ds.$$

El valor

$$u_p^* = \int_a^b \phi_p(s)u(s)\,ds$$

recibe el nombre de coeficiente de Fourier de $u(s)$ con respecto al sistema $\{\phi_p\}$.

Hilbert demuestra que se puede definir un sistema ortonormal completo para todo intervalo finito $[a,b]$, por ejemplo mediante el uso de polinomios. Demuestra una desigualdad generalizada de Bessel y finalmente que la condición

$$\int_a^b u^2(s)\,ds = \sum_{p=1}^{\infty} u_p^{*2}$$

es equivalente a la completitud.

Hilbert vuelve a considerar la ecuación integral

$$f(s) = \phi(s) + \int_a^b K(s, t)\phi(t)\,dt. \tag{35}$$

El núcleo $K(s,t)$, no necesariamente simétrico, puede desarrollarse en una serie de «Fourier» doble por medio de los coeficientes

$$a_{pq} = \int_a^b \int_a^b K(s, t)\phi_p(s)\phi_q(t)\, ds\, dt.$$

Se sigue que

$$\sum_{p,q=1}^{\infty} a_{pq}^2 \leq \int_a^b \int_a^b K^2(s, t)\, ds\, dt.$$

y también que si

$$a_p = \int_a^b \phi_p(s) f(s)\, ds,$$

es decir, si los a_p son los «coeficientes de Fourier» de $f(s)$, entonces $\Sigma_{p=1}^{\infty} a_p^2 < \infty$.

Hilbert transforma a continuación la anterior ecuación integral en un sistema de infinitas ecuaciones lineales en infinitas incógnitas. La idea consiste en tratar de resolver la ecuación integral en $\phi(s)$ como el problema de hallar los «coeficientes de Fourier» de $\phi(s)$. Representando los coeficientes, aún desconocidos, por x_1, x_2, ... obtiene las siguientes ecuaciones lineales

$$x_p + \sum_{q=1}^{\infty} a_{pq} x_q = a_p, \qquad p = 1, 2, \ldots \tag{36}$$

y demuestra que si este sistema tiene solución única entonces la ecuación integral tiene solución continua única, y que cuando el sistema lineal homogéneo asociado a (36) tiene n soluciones linealmente independientes, la ecuación integral homogénea asociada a (35)

$$0 = \phi(s) + \int_a^b K(s, t)\phi(t)\, dt, \tag{37}$$

tiene n soluciones linealmente independientes. En este caso la ecuación integral no homogénea original tiene una solución si y sólo si $\psi^{(h)}(s)$, $h = 1, 2, \ldots, n$, que son las n soluciones linealmente independientes de la ecuación homogénea transpuesta

$$\phi(s) + \int_a^b K(t, s)\phi(t)\, dt = 0$$

y que también existen cuando (37) tiene n soluciones, satisfacen las condiciones

$$\int_a^b \psi^{(h)}(s)f(s)\, ds = 0, \qquad h = 1, 2, ..., n. \tag{38}$$

Así se obtiene el teorema de la alternativa de Fredholm: o bien la ecuación

$$f(s) = \phi(s) + \int_a^b K(s, t)\phi(t)\, dt \tag{39}$$

tiene una única solución para toda función f, o la ecuación homogénea asociada tiene n soluciones linealmente independientes. En el segundo caso (39) tiene una solución si y sólo si se verifican las condiciones de ortogonalidad (38).

Hilbert considera a continuación el problema de autovalores

$$f(s) = \phi(s) - \lambda \int_a^b K(s, t)\phi(t)\, dt \tag{40}$$

donde K es ahora simétrico. La simetría de K implica que sus «coeficientes de Fourier» determinan una forma cuadrática $K(x,x)$ que es completamente continua, y demuestra que existe una transformación ortogonal T cuya matriz es $\{l_{pq}\}$ tal que

$$K(x', x') = \sum_{p=1}^{\infty} \mu_p x_p'^2,$$

donde los μ_p son los autovalores recíprocos de la forma cuadrática $K(x,x)$. Las autofunciones $\{\phi_p(s)\}$ para el núcleo $K(s,t)$ vienen definidas ahora por

$$L_p(K(s)) = \sum_{q=1}^{\infty} l_{pq} \int_a^b K(s, t)\Phi_q(t)\, dt = \mu_p \phi_p(s)$$

donde los $\Phi_q(t)$ constituyen un sistema ortonormal completo dado. Se puede demostrar que los $\phi_p(s)$ (distintos de los $\Phi_q(t)$ forman también un sistema ortonormal y satisfacen

$$\phi_p(s) = \lambda_p \int_a^b K(s, t)\phi_p(t)\, dt$$

donde $\lambda_p = 1/\mu_p$. Así demuestra Hilbert de nuevo la existencia de autofunciones para el caso homogéneo de (40) y para todo autovalor finito de la forma cuadrática $K(x,x)$ asociada al núcleo $K(s,t)$ de (40).

Hilbert demuestra de nuevo (teorema de Hilbert-Schmidt) que si $f(s)$ es una función continua cualquiera para la que existe una g tal que

$$\int_a^b K(s, t)g(t)\, dt = f(s),$$

entonces f es representable en serie de autofunciones de K, que es uniforme y absolutamente convergente (véase [31]). Hilbert utiliza este resultado para demostrar que la ecuación homogénea asociada a (40) no tiene ninguna solución no trivial excepto para los autovalores λ_p. Entonces el teorema de la alternativa de Fredholm afirma que: para $\lambda \neq \lambda_p$ la ecuación (40) tiene solución única; para $\lambda = \lambda_p$ tiene solución si y sólo si se satisfacen las n_p condiciones

$$\int_a^b \phi_{p+j}(s)f(s) = 0, \qquad j = 1, 2, ..., n_p$$

donde las $\phi_{p+j}(s)$ son las n_p autofunciones asociadas a λ_p. Por último demuestra una vez más la generalización del teorema de los ejes principales:

$$\int_a^b \int_a^b K(s, t)u(s)u(t)\, ds\, dt = \sum_{p=1}^{\infty} \frac{1}{\lambda_p} \left\{\int_a^b u(t)\phi_p(t)\, dt\right\}^2,$$

donde $u(s)$ es una función continua arbitraria y todas las ϕ_p asociadas a cualquier λ_p aparecen incluidas en la suma.

En este último trabajo (1906) Hilbert prescinde del uso de los determinantes infinitos de Fredholm, demostrando directamente la

relación entre las ecuaciones integrales y la teoría de sistemas ortogonales completos para el desarrollo en serie de funciones.

Hilbert aplicó también sus resultados sobre ecuaciones integrales a diversos problemas de geometría y de física. En particular, en el tercero de los seis artículos resolvió el problema de Riemann de construir una función holomorfa en un dominio limitado por una curva lisa, cuando se da la parte real o imaginaria del valor en la frontera o ambas están relacionadas por una ecuación lineal dada.

Uno de los resultados más notables de la obra de Hilbert, que apareció en los artículos de 1904 y 1905, es la formulación de los problemas de contorno de Sturm-Liouville para ecuaciones diferenciales como ecuaciones integrales. El resultado de Hilbert afirma que los autovalores y autofunciones de la ecuación diferencial

$$\frac{d}{dx}\left(p(x)\,\frac{du}{dx}\right) + q(x)u + \lambda u = 0 \tag{41}$$

con las condiciones de contorno

$$u(a) = 0, \qquad u(b) = 0$$

(e incluso condiciones de contorno más generales) son los autovalores y autofunciones de

$$\phi(x) - \lambda \int_a^b G(x, \xi)\phi(\xi)\, d\xi = 0, \tag{42}$$

donde $G(x,\xi)$ es la función de Green para la ecuación (41), es decir, una solución particular de

$$\frac{d}{dx}\left(p\,\frac{du}{dx}\right) + q(x)u = 0$$

que satisface ciertas condiciones de diferenciabilidad y cuya derivada parcial $\partial G/\partial x$ tiene una singularidad de salto en $x = \xi$ igual a $-1/p(\xi)$. Resultados análogos se verifican para las ecuaciones en derivadas parciales. Así pues, las ecuaciones integrales constituyen un instrumento para resolver ecuaciones diferenciales ordinarias y en derivadas parciales.

Recapitulando los principales resultados obtenidos por Hilbert,

en primer lugar estableció la teoría espectral general para núcleos simétricos *K*. Solamente veinte años antes había requerido grandes esfuerzos matemáticos (cap. 28, sec. 8) demostrar la existencia de la frecuencia de oscilación inferior para una membrana. Con ayuda de las ecuaciones integrales, la demostración constructiva de la existencia de la serie completa de frecuencias y de las autofunciones concretas se obtenía para condiciones muy generales relativas al medio de oscilación. Emile Picard [15] fue el primero en obtener estos resultados utilizando la teoría de Fredholm. Otro resultado notable debido a Hilbert es el de que el desarrollo de una función en las autofunciones pertenecientes a una ecuación integral de segundo tipo depende de la resolubilidad de la correspondiente ecuación integral de primer tipo. En particular, Hilbert descubrió que el éxito del método de Fredholm radicaba en el concepto de continuidad completa, que aplicó a las formas bilineales, estudiándolo sistemáticamente. Aquí inauguraba la teoría espectral de formas bilineales simétricas.

Una vez que Hilbert mostró cómo convertir problemas de ecuaciones diferenciales en ecuaciones integrales, se empezó a utilizar este planteamiento cada vez más para resolver problemas físicos; aquí el uso de una función de Green para la conversión ha resultado una herramienta de gran importancia. Hilbert mismo demostró,[16] por su parte, en problemas de dinámica de gases, que uno puede ir directamente a las ecuaciones integrales. Este recurso directo a las ecuaciones integrales es posible porque el concepto de suma demuestra ser tan fundamental en algunos problemas físicos como el concepto de cociente incremental que conduce a las ecuaciones diferenciales en otros. Hilbert subrayó también que no eran las ecuaciones diferenciales ordinarias ni en derivadas parciales, sino las ecuaciones integrales, el punto de partida natural y necesario para la teoría de desarrollos de funciones en serie, y que los desarrollos obtenidos mediante ecuaciones diferenciales no eran más que casos particulares del teorema general de la teoría de ecuaciones integrales.

---

[15] Rendiconti del Circolo Matematico di Palermo, 22, 1906, 241-259.

[16] Math. Ann., 72, 1912, 562-577 = *Grundzuge*, cap. 22.

4. Los sucesores inmediatos de Hilbert

La obra de Hilbert sobre ecuaciones integrales fue simplificada por Erhard Schmidt (1876-1959), profesor en varias universidades alemanas, utilizando métodos que había introducido H. A. Schwarz en teoría del potencial. La contribución más importante de Schmidt fue su generalización del concepto de autofunción para ecuaciones integrales con núcleo no simétrico, que data de 1907.[17]

Otro matemático que también se dedicó, en 1907, a continuar la obra de Hilbert fue Friedrich Riesz (1880-1956), de nacionalidad húngara.[18] Hilbert había estudiado ecuaciones integrales de la forma

$$f(s) = \phi(s) + \int_a^b K(s, t)\phi(t)\, dt,$$

donde f y K son continuas. Riesz trató de extender las ideas de Hilbert a funciones $f(s)$ más generales. Con este objeto, lo que se necesitaba era asegurarse de que los «coeficientes de Fourier» de f con respecto a un sistema ortonormal de funciones $\{\phi_p\}$ se podían calcular. También estaba interesado en investigar en qué condiciones una sucesión de números dada $\{a_p\}$ podía ser la sucesión de coeficientes de Fourier de una función f con respecto a un sistema ortonormal dado $\{\phi_p\}$.

Riesz consideró funciones cuyo cuadrado es integrable en el sentido de Lebesgue, y obtuvo el siguiente teorema: sea $\{\phi_p\}$ una sucesión ortonormal de funciones de cuadrado integrable en el sentido de Lebesgue, definidas sobre el intervalo $[a,b]$. Si $\{a_p\}$ es una sucesión de números reales, entonces la convergencia de $\sum_{p=1}^{\infty} a_p^2$ es condición necesaria y suficiente para que exista una función f tal que

$$\int_a^b f(x)\phi_p(x)\, dx = a_p$$

para cada ϕ_p y a_p. Se demuestra además que la función f es de cuadrado integrable. Este teorema establece una correspondencia biunívoca entre el conjunto de las funciones de cuadrado integrable y el

[17] Math. Ann., 63, 1907, 433-476 y 64, 1907, 161-174.
[18] Comp. Rend., 144, 1907, 613-619, 734-736, 1409-1411.

conjunto de las sucesiones de cuadrado sumable, para cada sucesión ortonormal de funciones de cuadrado integrable.

Utilizando las funciones integrables en el sentido de Lebesgue pudo demostrar Riesz que la ecuación integral de segundo tipo

$$f(s) = \phi(s) + \int_a^b K(s, t)\phi(t)\, dt,$$

admite solución bajo las condiciones más débiles de que $f(s)$ y $K(s,t)$ sean de cuadrado integrable. La solución es única salvo una función cuya integral de Lebesgue sobre $[a,b]$ sea nula.

El mismo año que Riesz publicaba sus primeros artículos, un profesor de la universidad de Colonia, Ernst Fischer (1875-1959) introducía el concepto de convergencia en media.[19] Se dice que una sucesión de funciones $\{f_n\}$ definidas sobre el intervalo $[a,b]$ converge en media si

$$\lim_{m,n\to\infty} \int_a^b (f_n(x) - f_m(x))^2\, dx = 0$$

y se dice que $\{f_n\}$ converge en media a f si

$$\lim_{n\to\infty} \int_a^b (f - f_n)^2\, dx = 0$$

donde las integrales están tomadas en el sentido de Lebesgue. La función f está determinada de manera única salvo una función definida sobre un conjunto de medida cero, es decir, una función $g(x) \neq 0$, llamada una función nula, que satisface la condición $\int_a^b g^2(x)\, dx = 0$.

El conjunto de funciones de cuadrado integrable en el sentido de Lebesgue sobre un intervalo $[a,b]$ se representó más tarde por $L^2(a,b)$ o simplemente por L^2. El resultado principal obtenido por Fischer fue que $L^2(a,b)$ es completo en media, es decir, si las funciones f_n pertenecen a $L^2(a,b)$ y la sucesión $\{f_n\}$ converge en media, entonces converge en media a una función f perteneciente a $L^2(a,b)$. Esta propiedad de completitud es la principal ventaja de utilizar funciones de cuadrado sumable. Fischer dedujo a continuación como corolario el teorema de Riesz mencionado anteriormente, y que hoy se conoce

---

[19] Comp. Rend., 144, 1907, 1022-1024.

como teorema de Riesz-Fischer. En una nota posterior [20] subraya Fischer que era esencial el uso de las funciones de cuadrado integrable; el teorema no sería válido para ningún subconjunto.

La determinación de una función $f(x)$ correspondiente a un conjunto de coeficientes de Fourier $\{a_n\}$ con respecto a una sucesión dada de funciones ortonormales $\{g_n\}$, o la determinación de una f tal que

$$\int_a^b g_n(x)f(x)\,dx = a_n, \qquad n = 1, 2, \ldots,$$

que aparecía en los artículos de Riesz de 1907, recibe hoy el nombre de problema de los momentos (entendiendo siempre la integración en el sentido de Lebesgue). En 1910 [21] intentó Riesz generalizar este problema, y debido a que necesitaba utilizar las desigualdades de Hölder

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |b_i|^q\right)^{1/q}$$

y

$$\left|\int_M f(x)g(x)\,dx\right| \leq \left(\int_M |f|^p\,dx\right)^{1/p}\left(\int_M |g|^q\,dx\right)^{1/q},$$

donde $1/p + 1/q = 1$, y otras desigualdades, se vio obligado a introducir el conjunto L^p de las funciones f medibles sobre un conjunto M tales que $|f|^p$ es integrable sobre M. Su primer teorema importante fue el de que si una función $h(x)$ es tal que el producto $f(x)h(x)$ es integrable para toda f de L^p, entonces h pertenece a L^q y, recíprocamente, el producto de una función de L^p por otra de L^q siempre es integrable; se entiende que en todos los casos es $p > 1$ y $1/p + 1/q = 1$.

Riesz introdujo también los conceptos de convergencia fuerte y débil. La sucesión de funciones $\{f_n\}$ converge fuertemente a f (en la media de orden p) si

---

[20] Comp. Rend., 144, 1907, 1148-1150.

[21] Math. Ann., 69, 1910, 449-497.

$$\lim_{n\to\infty} \int_a^b |(f_n(x) - f(x))|^p \, dx = 0$$

La sucesión $\{f_n\}$ converge débilmente a f si

$$\int_a^b |f_n(x)|^p \, dx < M$$

donde M es independiente de n, y si para todo x del intervalo $[a,b]$

$$\lim_{n\to\infty} \int_a^x (f_n(t) - f(t)) \, dt = 0$$

La convergencia fuerte implica la convergencia débil. (La definición moderna de convergencia débil: si las funciones f_n y f pertenecen a L^p y si

$$\lim_{n\to\infty} \int_a^x (f(x) - f_n(x)) \, g(x) \, dx = 0$$

se verifica para toda g de L^q, entonces $\{f_n\}$ converge débilmente a f, es equivalente a la de Riesz).

En el mismo artículo de 1910 generaliza Riesz la teoría de las ecuaciones integrales

$$\phi(x) - \lambda \int_a^b K(x, t)\phi(t) \, dt = f(x)$$

al caso en que la f dada y la incógnita ϕ son funciones de L^p. Los resultados sobre la resolución del problema de los autovalores para esta ecuación integral son análogos a los de Hilbert. Lo más sorprendente es que, para llevar esto a cabo, Riesz introduce el concepto abstracto de operador, formula para él el concepto de continuidad completa de Hilbert, inaugurando así la teoría abstracta de operadores. En el próximo capítulo hablaremos más de este planteamiento abstracto. Entre otros resultados demostró Riesz que el espectro continuo de un operador real completamente continuo sobre L^2 es vacío.

5. Generalizaciones de la teoría

La importancia que atribuyó Hilbert a las ecuaciones integrales puso de moda el tema a escala mundial durante un largo período de tiempo, durante el cual se produjo una masa enorme de literatura, la mayor parte de un valor efímero. Sin embargo, algunas generalizaciones se mostraron valiosas, aunque aquí solamente podremos mencionarlas.

La teoría de ecuaciones integrales que hemos presentado se refiere a ecuaciones integrales lineales, es decir, en las que la función incógnita $u(x)$ aparece linealmente. Esta teoría se ha generalizado a ecuaciones integrales no lineales donde la función incógnita aparece elevada al cuadrado o potencias más altas, o de alguna manera más complicada.

Por otra parte, en nuestro breve resumen apenas hemos dicho nada acerca de las condiciones sobre las funciones $f(x)$ y $K(x,\xi)$ que daban validez a las conclusiones. Si estas funciones no son continuas o si el intervalo $[a,b]$ se sustituye por un intervalo infinito, muchos de los resultados se ven afectados, o al menos se necesitan nuevas demostraciones. Así, incluso la transformación de Fourier

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos (x\xi) u(\xi)\, d\xi,$$

que puede considerarse como una ecuación integral de primer tipo y que tiene como solución la transformación inversa

$$u(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos (x\xi) f(\xi)\, d\xi,$$

tiene exactamente dos autovalores ± 1 y cada uno de ellos tiene una cantidad infinita de autofunciones. Estos casos se estudian ahora bajo el nombre de ecuaciones integrales singulares; tales ecuaciones no pueden resolverse por los métodos aplicados a las ecuaciones de Volterra y Fredholm. Además presentan una curiosa propiedad, a saber, que hay intervalos continuos de valores de λ o espectros de banda para los que existen soluciones. El primer trabajo importante sobre este tema se debe a Hermann Weyl (1885-1955).[22]

---

[22] Math. Ann., 66, 1908, 273-324 = *Ges. Abh.*, 1, 1-86.

También han recibido mucha atención los teoremas de existencia para ecuaciones integrales, tanto lineales como no lineales. Por ejemplo, muchos matemáticos han formulado teoremas de existencia para ecuaciones como

$$y(x) = f(x) + \int_a^x K(x, s, y(s))\, ds,$$

que incluyen como caso especial la ecuación de Volterra de segundo tipo

$$y(x) = f(x) + \int_a^x K(x, s)y(s)\, ds.$$

Desde un punto de vista histórico, el siguiente desarrollo importante consistió en la generalización de las ideas anteriores en un marco abstracto. Hilbert consideraba dada una función por sus coeficientes de Fourier, los cuales satisfacían la condición de que $\Sigma_1^\infty a_p^2$ es finito, por lo que introdujo también sucesiones generales de números reales $\{x_n\}$, tales que $\Sigma_1^\infty x_n^2$ es finito. Riesz y Fischer demostraron entonces que hay una correspondencia biunívoca entre las funciones de cuadrado integrable y las sucesiones de cuadrado sumable de sus coeficientes de Fourier. Las sucesiones de cuadrado sumable pueden ser consideradas como puntos de un espacio de infinitas dimensiones que es una generalización del espacio euclídeo n-dimensional. Así pues, las funciones pueden ser consideradas como puntos de un espacio, llamado hoy espacio de Hilbert, y la integral $\int_a^b K(x,y)\, u(x)\, dx$ puede ser interpretada como un operador que transforma $u(x)$ en sí misma o en otra función. Estas ideas sugirieron un planteamiento abstracto para el estudio de las ecuaciones integrales que encajaba bien en el también incipiente planteamiento abstracto del cálculo de variaciones. Este nuevo planteamiento se conoce hoy como análisis funcional y a él dedicaremos el próximo capítulo.

Bibliografía

Bernkopf, M.: «The Development of Function Spaces with Particular Reference to their Origins in Integral Equation Theory». *Archive for History of Exact Sciences,* 3, 1966, 1-96.

Bliss, G. A.: «The Scientific Work of E. H. Moore». *Amer. Math. Bull*, 40, 1934, 501-514.

Bocher, M.: *An Introduction to the Study of Integral Equations*, 2.ª ed., Cambridge University Press, 1913.

Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.

Davis, Harold T.: *The Present State of Integral Equations*, Indiana University Press, 1926.

Hahn, H.: «Berich über die Theorie der linearen Integralgleichungen». *Jahres. der Deut. Math.-Verein.*, 20, 1911, 69-117.

Hellinger, E.: *Hilberts Arbeiten über Integralgleichungen und unendliche Gleichungs-systeme*, in Hilbert's *Gesam. Abh.*, 3, 94-145, Julius Springer, 1935.

— «Begründung der Theorie quadratischer Formen von unendlichvielen Veränderliche». *Jour. für Math.*, 136, 1909, 210-271.

Hellinger, E., y Toeplitz, O.: «Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, vol. 2, part. 3, 2.ª half, 1335-1597.

Hilbert, D.: *Grundzüge einer állgemeinen Theorie der linearen Integralgleichungen*, 1912, Chelsea (reprint), 1953.

Reid, Constance: *Hilbert*, Springer-Verlag, 1970.

Volterra, Vito: *Opere matematiche*, 5 vols. Accademia Nazionale dei Lincei, 1954-1962.

Weyl, Hermann: *Gesammelte Abhandlungen*, 4 vols., Springer-Verlag, 1968.

Capítulo 46
EL ANALISIS FUNCIONAL

> Uno debe estar bien seguro de que ha permitido a la ciencia hacer un gran progreso, si va a sobrecargarla con una multitud de términos nuevos y a exigir que los lectores sigan una investigación que les ofrece tantas cosas extrañas.
>
> A. L. CAUCHY

1. ¿Qué es el análisis funcional?

A finales del siglo XIX se puso claramente de manifiesto que muchos campos de la matemática utilizaban transformaciones u operadores que actuaban sobre funciones; por ejemplo, incluso la simple operación de diferenciación ordinaria y su inversa la antidiferenciación o integración actúan sobre una función para producir otra. En los problemas de cálculo de variaciones en los que se manejan integrales tales como

$$J = \int_a^b F(x, y, y')\, dx$$

dicha integral puede considerarse como operando sobre una clase de funciones $y(x)$, de las cuales se busca precisamente aquella que haga máxima o mínima la integral. La teoría de ecuaciones diferenciales nos ofrece otra clase de operadores; por ejemplo, el operador diferencial

$$L = \frac{d^2}{dx^2} + p(x)\frac{d}{dx} + q(x)$$

actuando sobre una clase de funciones $y(x)$ las transforma en otras funciones. Naturalmente, resolver la ecuación diferencial asociada a este operador significa buscar una función particular $y(x)$ tal que L actuando sobre ella nos dé 0, y que satisfaga posibles condiciones iniciales o de contorno dadas. Como un último ejemplo de operadores tenemos las ecuaciones integrales; el segundo miembro de la igualdad

$$f(x) = \int_a^b K(x, y)u(x)\, dx$$

puede ser considerado como un operador que actúa sobre las diversas $u(x)$ posibles para producir nuevas funciones, y aquí también, como en el caso de las ecuaciones diferenciales, la función $u(x)$ solución de la ecuación debe transformarse en $f(x)$.

La idea que motivó la creación del análisis funcional fue la de que todos estos operadores podrían ser estudiados dentro de una formulación abstracta de una teoría general de operadores actuando sobre clases de funciones; además, estas funciones podrían ser consideradas como elementos o puntos de un espacio, entonces el operador transformaría puntos en puntos, y en este sentido sería una generalización de las transformaciones geométricas ordinarias, tales como las rotaciones. Algunos de los operadores anteriores transforman funciones en números reales y no en otras funciones; los operadores que dan como resultado números reales o complejos reciben actualmente el nombre de funcionales, mientras que el nombre de operador se reserva normalmente para las transformaciones que aplican funciones en funciones. Así pues, el nombre de análisis funcional, introducido por Paul P. Lévy (1886) cuando los funcionales constituían el centro del interés, ya no resulta hoy muy adecuado. La búsqueda de la generalidad y de la unificación de campos diversos es una de las características distintivas de la matemática del siglo XX, y el análisis funcional también tiende a alcanzar esas metas, evidentemente.

2. La teoría de funcionales

La teoría abstracta de funcionales fue iniciada por Volterra en sus trabajos sobre cálculo de variaciones. Volterra publicó una serie

de artículos [1] sobre funciones de líneas (curvas), tal como él las llamaba. Una función de línea era, para Volterra, una función real F, cuyos valores dependen de todos los valores que toman ciertas funciones $y(x)$ definidas en un intervalo $[a,b]$. Las funciones mismas se consideraban como puntos de un espacio en el cual se podían definir los entornos de un punto y el límite de una sucesión de puntos. Volterra dio, para los funcionales $F[y(x)]$, definiciones de continuidad, derivada y diferencial, pero estas definiciones no resultaban adecuadas para la teoría abstracta del cálculo de variaciones y fueron reemplazadas; de hecho, las definiciones dadas por Volterra fueron criticadas por Hadamard.[2]

Antes incluso de que Volterra comenzara sus trabajos, ya había sido admitida la idea de considerar una colección de funciones $y(x)$, definidas todas ellas en algún intervalo común, como puntos de un espacio. Riemann hablaba ya en su tesis de una colección de funciones formando un dominio cerrado conexo (de puntos de un espacio), y Giulio Ascoli (1843-1896) [4] y Cesare Arzelà [5] intentaban, por su parte, extender a conjuntos de funciones la teoría de conjuntos de puntos de Cantor, considerando así las funciones de nuevo como puntos de un espacio; Arzelà hablaba también de funciones de línea. Hadamard sugería en el Primer Congreso Internacional de Matemáticos, en 1897,[6] considerar las curvas como puntos de un conjunto; él pensaba en la familia de todas las funciones continuas definidas en el intervalo [0,1], familia que se le presentó de manera natural en sus trabajos sobre ecuaciones en derivadas parciales. Emile Borel hizo la misma sugerencia,[7] aunque con una finalidad diferente, la del estudio de funciones arbitrarias por medio de series.

Hadamard también se vio conducido al estudio de los funcionales[8] con motivo del cálculo de variaciones. El nombre de funcional es debido a él, así como el llamar a un funcional $U[y(t)]$ lineal cuan-

---

[1] *Atti della Reale Accademia dei Lincei*, (4), 3, 1887, 97-105, 141-146, 153-158 = *Opere Matematiche*, 1, 294-314, y otros del mismo año y posteriores.

[2] Bull. Soc. Math. de France, 30, 1902, 40-43 = *Œuvres*, 1, 401-404.

[3] *Werke*, p. 30.

[4] Memorie della Reale Accademia dei Lincei, (3), 18, 1883, 521-586.

[5] Atti della Accademia dei Lincei, Rendiconti, (4), 5, 1889, 342-348.

[6] *Verhandlungen des ersten internationalen Mathematiker-Kongresses*, Teubner, 1898, 201-202.

[7] Verhandlungen, 204-205.

[8] Comp. Rend., 136, 1903, 351-354 = *Œuvres*, 1, 405-408.

do, si $y(t) = \lambda_1 y_1(t) + \lambda_2 y_2(t)$ donde λ_1 y λ_2 son constantes, entonces $U[y(t)] = \lambda_1 U[y_1(t)] + \lambda_2 U[y_2(t)]$.

El primer esfuerzo importante por construir una teoría abstracta de espacios de funciones y de funcionales fue el llevado a cabo por Maurice Fréchet (1878), eminente matemático francés, en su tesis doctoral de 1906.[9] En lo que Fréchet llamó cálculo funcional, intentó unificar en términos abstractos las ideas contenidas en los trabajos de Cantor, Volterra, Arzelà, Hadamard y otros.

Con objeto de conseguir el más alto grado de generalidad para sus espacios de funciones, Fréchet echó mano de todas las ideas básicas sobre conjuntos desarrolladas por Cantor, aunque para él los puntos de los conjuntos eran ahora funciones. También formuló de una manera más general el concepto de límite de un conjunto de puntos; este concepto no quedaba definido explícitamente, sino que venía caracterizado por unas propiedades lo bastante generales como para incluir los distintos tipos de límites que aparecían en las teorías concretas que Fréchet intentaba unificar. Introdujo así una clase L de espacios, donde la L indica que existe un concepto de límite definido sobre cada uno de estos espacios, de manera que si B es un espacio de la clase L y tomamos en él elementos A_1, A_2, ... arbitrarios, debe ser posible determinar si existe o no un elemento único A, llamado, cuando existe, límite de la sucesión $\{A_n\}$, tal que

a) Si $A_i = A$ para todo i, entonces $\lim\{A_n\} = A$.

b) Si A es el límite de $\{A_n\}$ entonces A también es el límite de toda subsucesión infinita de $\{A_n\}$.

A partir de este concepto definía Fréchet un cierto número de conceptos aplicables a cualquier espacio de la clase L; por ejemplo, el conjunto derivado E' de un conjunto E es el conjunto de todos los puntos del espacio en cuestión que son límites de sucesiones de puntos de E, E es cerrado si E' está contenido en E; E es perfecto si $E' = E$; un punto A perteneciente a E es un punto interior de E (en sentido restringido) si A no es límite de ninguna sucesión contenida en el complementario de E; un conjunto E es compacto si es finito o si todo subconjunto infinito de E tiene por lo menos un elemento límite; si E es cerrado y compacto recibe el nombre de extremal («compacto» según Fréchet es lo que actualmente se llama relativamente secuencialmente compacto, y «extremal» equivale al concepto actual de secuencialmente compacto). El primer teorema

[9] Rendiconti del Circolo Matematico di Palermo, 22, 1906, 1-74.

importante de Fréchet es una generalización del teorema de los intervalos cerrados encajados: si $\{E_n\}$ es una sucesión decreciente de subconjuntos cerrados de un conjunto extremal, es decir, tal que E_{n+1} está contenido en E_n para todo n, entonces la intersección de todos los E_n es no vacía.

Fréchet pasa a considerar entonces funcionales (a los que llama operaciones funcionales), es decir, funciones con valores reales definidas sobre un conjunto E, y define la continuidad de un funcional de la manera siguiente: un funcional U se llamará continuo en un elemento A de E, si se verifica que lím $U(A_n) = U(A)$ para toda sucesión $\{A_n\}$ contenida en E y que converja a A. Introduce también la idea de semicontinuidad de funcionales, concepto definido ya para funciones ordinarias por René Baire (1874-1932) en 1899.[10] El funcional U se llamará semicontinuo superiormente en E si $U(A) \geqslant$ lím sup $U(A_n)$ para las $\{A_n\}$ como en el caso anterior, y será semicontinuo inferiormente si $U(A) \geqslant$ lím inf $U(A_n)$.[11]

A partir de estas definiciones pudo demostrar Fréchet una serie de teoremas sobre funcionales. Por ejemplo, todo funcional continuo sobre un conjunto extremal E está acotado y alcanza un valor máximo y un valor mínimo sobre E; todo funcional semicontinuo superiormente definido sobre un conjunto extremal E, está acotado superiormente y alcanza su máximo sobre E.

A continuación Fréchet introdujo generalizaciones de conceptos aplicables a sucesiones y conjuntos de funcionales, tales como el de convergencia uniforme, convergencia cuasi-uniforme, compacidad y equicontinuidad. Por ejemplo, la sucesión de funcionales $\{U_n\}$ converge uniformemente a U si, dado cualquier número positivo ε, $|U_n(A) - U(A)| < \varepsilon$ tomando n suficientemente grande pero independiente del punto A de E. De esta manera pudo demostrar para el caso de funcionales generalizaciones de teoremas obtenidos previamente para funciones reales.

Una vez estudiados los espacios generales L, definió Fréchet otros espacios más especializados, tales como los espacios de entornos, redefinió los conceptos utilizados para espacios con puntos límites, y demostró teoremas análogos a los anteriores, pero a menudo con mejores resultados, debido a que los espacios eran más ricos en propiedades.

---

[10] Annali di Mat., (3), 13, 1899, 1-122.

[11] El límite inferior es el mínimo punto límite de la sucesión $U(A_n)$.

Por último introdujo Fréchet los espacios métricos. En un espacio métrico está definida una función que juega el papel de la distancia *(écart)*, representada por (A,B) para cada par de puntos A y B del espacio, función que debe satisfacer las siguientes condiciones:

a) $(A,B) = (B,A) \geqslant 0$.

b) $(A,B) = 0$ si y sólo si $A = B$.

c) $(A,B) + (B,C) \geqslant (A,C)$.

La condición (c) se llama desigualdad triangular. A estos espacios los llamó Fréchet de clase ε, y también pudo demostrar para ellos un cierto número de teoremas sobre sus funcionales, de una manera muy parecida a los casos de los espacios más generales.

Fréchet dio algunos ejemplos de espacios de funciones: por ejemplo, el conjunto de todas las funciones reales de una variable real continuas sobre un intervalo I, con el «*écart*» entre dos funciones cualesquiera f y g definido de la forma $\max_{x \in I} |f(x) - g(x)|$, es un espacio de clase ε; hoy día a este «*écart*» se le denomina norma del máximo.

Otro ejemplo propuesto por Fréchet es el del conjunto de todas las sucesiones de números reales; si $x = (x_1, x_2, \ldots)$ e $y = (y_1, y_2, \ldots)$ son dos de estas sucesiones, el «*écart*» entre x e y se define como

$$(x, y) = \sum_{p=1}^{\infty} \frac{1}{p!} \frac{|x_p - y_p|}{1 + |x_p - y_p|}$$

obteniéndose de esta forma un espacio de dimensión infinita numerable, como Fréchet mismo hizo notar.

Usando sus espacios ε, consiguió dar Fréchet [12] una definición general de continuidad, diferencial y diferenciabilidad de un funcional. Aunque estas definiciones no eran del todo adecuadas para el cálculo de variaciones, su definición de diferencial vale la pena destacarla porque constituye el núcleo de lo que demostró ser más satisfactorio. Fréchet supone la existencia de un funcional lineal $L(\eta(x))$ tal que

$$F[y(x) + \eta(x)] = F(y) + L(\eta) + \varepsilon M(\eta)$$

donde $\eta(x)$ es una variación sobre la $y(x)$, $M(\eta)$ es el máximo del

[12] Amer. Math. Soc. Trans., 15, 1914, 135-161.

valor absoluto de $\eta(x)$ sobre $[a,b]$ y ε tiende a cero con M; entonces $L(\eta)$ es, por definición, la diferencial de $F(y)$. Fréchet presuponía también la continuidad de $F(y)$, lo cual es más de lo que puede asegurarse normalmente en los problemas de cálculo de variaciones.

Charles Albert Fischer (1884-1922) [13] mejoró posteriormente la definición de Volterra de la derivada de un funcional, de manera que cubriese el caso de los funcionales usados en el cálculo de variaciones; la diferencial de un funcional podía definirse entonces en términos de la derivada.

Por lo que se refiere a las definiciones básicas de las propiedades de los funcionales necesarias en el cálculo de variaciones, las formulaciones finales fueron dadas por Elizabeth le Stourgeon (1881-1971).[14] El concepto clave, el de diferencial de un funcional, es una modificación del de Fréchet: se dice que el funcional $F(y)$ tiene una diferencial en $y_0(x)$ si existe un funcional lineal $L(\eta)$ tal que para todos los arcos $y_0 + \eta$ en un entorno de y_0, se verifica la relación

$$F(y_0 + \eta) = F(y_0) + L(\eta) + M(\eta) \cdot \varepsilon(\eta)$$

donde $M(\eta)$ es el máximo del valor absoluto de η y η' sobre el intervalo $[a,b]$ y $\varepsilon(\eta)$ tiende a cero con $M(\eta)$. También definió las diferenciales segundas.

Tanto Le Stourgeon como Fischer obtuvieron, a partir de sus definiciones de las diferenciales, condiciones necesarias para que un funcional admita un mínimo, que esta vez sí eran aplicables a los problemas del cálculo de variaciones. Por ejemplo, una condición necesaria para un funcional $F(y)$ tenga un mínimo para $y = y_0$ es que $L(\eta)$ se anule para toda $\eta(x)$ que sea continua y tenga derivada primera continua en $[a,b]$ y tal que $\eta(a) = \eta(b) = 0$. Se puede deducir la ecuación de Euler de la condición de que se anule la diferencial primera, y usando una definición adecuada de la diferencial segunda de un funcional (que ya habían dado varios de los matemáticos que hemos mencionado), es posible deducir la necesidad de la condición de Jacobi del cálculo de variaciones.

El trabajo definitivo, por lo menos hasta 1925, en la teoría de funcionales necesaria para el cálculo de variaciones, fue llevado a

---

[13] Amer. J. Math., 33, 1913, 369-394.

[14] Amer. Math. Soc. Trans., 21, 1920, 357-383.

cabo por Leonida Tonelli (1885-1946), profesor de las universidades de Bolonia y Pisa. Después de haber escrito una gran cantidad de artículos sobre el tema desde 1911, publicó su «*Fondamenti di calcolo delle variazioni*» (2 vols., 1922, 1924), donde enfoca el tema desde el punto de vista de los funcionales. La teoría clásica se basaba principalmente en la teoría de ecuaciones diferenciales, y la intención de Tonelli era la de reemplazar los teoremas de existencia para ecuaciones diferenciales por teoremas de existencia para minimizar integrales de curvas. A lo largo de su obra el concepto de semicontinuidad inferior de un funcional es el concepto fundamental, porque los funcionales no suelen ser continuos.

Tonelli considera en primer lugar conjuntos de curvas y da teoremas que aseguran la existencia de una curva límite de una cierta clase de curvas. Los teoremas siguientes aseguran que la integral usual, pero en la forma paramétrica

$$\int_{t_1}^{t_2} F(t, x(t), y(t), x', y')\, dt$$

será semicontinua inferiormente como función de $x(t)$ y de $y(t)$ (posteriormente considera las integrales no paramétricas, más fundamentales), y obtiene las cuatro condiciones necesarias clásicas del cálculo de variaciones para el tipo habitual de problemas. El segundo tomo de la obra está dedicado principalmente a teoremas de existencia para una gran variedad de problemas, obtenidos a partir del concepto de semicontinuidad. Es decir, dada una integral de la forma anterior, demuestra, imponiendo condiciones sobre ella en tanto que funcional, y sobre la clase de curvas a considerar, que existe una curva en dicha clase que minimiza la integral. Sus teoremas se refieren a máximos y mínimos absolutos y relativos.

En cierta medida, la obra de Tonelli rinde también beneficios en el terreno de las ecuaciones diferenciales, dado que sus teoremas de existencia implican la existencia de soluciones de las ecuaciones diferenciales que suministraban las curvas minimales como soluciones en el planteamiento clásico. Sin embargo, su trabajo estaba naturalmente limitado a los tipos básicos de problemas del cálculo de variaciones. Aunque este enfoque abstracto se vio continuado posteriormente por muchos otros matemáticos, lo cierto es que no se han hecho grandes progresos en la aplicación de la teoría de funcionales al cálculo de variaciones.

3. El análisis funcional lineal

La labor más importante llevada a cabo en análisis funcional fue la de intentar dar una teoría abstracta para las ecuaciones integrales, en oposición al cálculo de variaciones. Las propiedades de los funcionales necesarias en el segundo caso son bastante particulares y no se suelen verificar para los funcionales en general. Por otro lado, el carácter no lineal de estos funcionales creaba dificultades que no se presentaban en el caso de los funcionales y operadores que incluyen como caso particular a las ecuaciones integrales. Al mismo tiempo que se hacían progresos concretos en la teoría de ecuaciones integrales, debidos a Schmidt, Fischer y Riesz, estos mismos matemáticos, y otros, comenzaron a trabajar en la elaboración de la correspondiente teoría abstracta.

La primera tentativa de elaborar una teoría abstracta de funcionales y operadores lineales fue hecha por el matemático americano E. H. Moore a partir de 1906.[15] Moore comprobó que había ciertas características comunes entre la teoría de ecuaciones lineales con un número finito de incógnitas, la teoría de sistemas infinitos de ecuaciones con un número infinito de incógnitas y la teoría de ecuaciones integrales lineales. Basándose en estas analogías, emprendió la tarea de construir una teoría abstracta, a la que llamó Análisis General, que incluiría a las teorías concretas anteriores como casos particulares, y adoptó para ello un planteamiento axiomático. No vamos a presentar aquí los detalles porque la influencia de Moore no fue muy extensa ni consiguió una metodología realmente eficaz; además su lenguaje simbólico era complicado y difícil de seguir.

El primer paso importante hacia una teoría abstracta de funcionales y operadores lineales fue dado por Erhard Schmidt[16] y Fréchet[17] en 1907. Hilbert, en sus trabajos sobre ecuaciones integrales, consideraba una función como dada por sus coeficientes de Fourier en su desarrollo con respecto a una sucesión ortonormal de funciones. Estos coeficientes, y los valores que asignaba a los x_i en su teoría de formas cuadráticas en infinitas variables, son sucesiones $\{x_n\}$ tales

[15] Véase, por ejemplo, *Atti del IV Congresso Internazionale dei Matematici* (1908), 2, Reale Accademia dei Lincei,, 1909, 98-114, y Amer. Math. Soc. Bull., 18, 1911-1912, 334-362.

[16] Rendiconti del Circolo Matematico di Palermo, 25, 1908, 53-77.

[17] Nouvelles Annales de Mathématiques, 8, 1908, 97-116, 289-317.

que $\Sigma_1^\infty x_n^2$ es finita. Sin embargo, Hilbert no llegó a considerar estas sucesiones como coordenadas de un punto en un espacio, ni utilizó tampoco un lenguaje geométrico. Este paso fue dado por Schmidt y Fréchet; considerando cada sucesión $\{x_n\}$ como un punto, las funciones quedaban representadas como puntos de un espacio de dimensión infinita. Schmidt consideraba también números complejos, además de reales, en las sucesiones $\{x_n\}$; un espacio de este tipo ha sido llamado desde entonces un espacio de Hilbert. Nosotros seguiremos en nuestra exposición el trabajo de Schmidt.

Los elementos de los espacios de funciones de Schmidt son sucesiones infinitas de números complejos, $z = \{z_p\}$ tales que

$$\sum_{p=1}^{\infty} |z_p|^2 < \infty.$$

Schmidt introdujo la notación $\|z\|$ para $\{\Sigma_{p=1}^\infty Z_p \cdot \overline{z}_p\}^{1/2}$, que recibió más tarde el nombre de norma de z. Siguiendo a Hilbert, Schmidt usó la notación (z,w) para $\Sigma_{p=1}^\infty z_p w_p$ con lo que $\|z\| = \sqrt{(z, \overline{z})}$ (modernamente se acostumbra definir directamente (z,w) como $\Sigma\Sigma_{p=1}^\infty z_p \overline{w}_p$). Dos elementos z y w del espacio se llaman ortogonales si y sólo si $(z,\overline{w}) = 0$. Schmidt demostró entonces una forma generalizada del teorema de Pitágoras, que asegura que si $z_1, z_2, \ldots, z_n$ son n elementos del espacio ortogonales dos a dos, entonces

$$w = \sum_{p=1}^{n} z_p$$

implica que

$$\|w\|^2 = \sum_{p=1}^{n} \|z_p\|^2$$

de lo cual se sigue que los n elementos del espacio ortogonales dos a dos son linealmente independientes. También obtuvo Schmidt la desigualdad de Bessel para este espacio general, en la forma: si $\{z_n\}$ es una sucesión ortonormal de elementos, de manera que $(z_p, \bar{z}_q) = \delta_{pq}$, y si w es un elemento cualquiera del espacio, entonces

$$\sum_{p=1}^{\infty} |(w, \overline{z}_p)|^2 \leq \|w\|^2.$$

También puede demostrarse la desigualdad de Schwarz y la desigualdad triangular para la norma.

Se dice que una sucesión de elementos $\{z_n\}$ converge fuertemente a z si $\|z_n - z\|$ tiende a cero, y entonces puede demostrarse que toda sucesión «fuerte» de Cauchy, es decir, toda sucesión para la cual $\|z_p - z_q\|$ tienda a cero cuando p y q tienden a ∞, converge a un cierto elemento z, de forma que el espacio de las sucesiones es completo, lo cual constituye una propiedad de importancia vital.

Schmidt introdujo a continuación el concepto de subespacio (fuertemente) cerrado: un subconjunto A del espacio considerado H se llama subespacio cerrado si es un subconjunto cerrado en el sentido de la convergencia que acabamos de definir, y si además es cerrado algebraicamente, es decir, si dados dos elementos w_1 y w_2 de A entonces también $a_1w_1 + a_2w_2$ pertenece a A, siendo a_1 y a_2 dos números complejos cualesquiera. Se prueba la existencia de tales subespacios cerrados simplemente tomando cualquier sucesión $\{z_n\}$ de elementos linealmente independientes y formando todas las combinaciones lineales finitas de elementos tomados de entre los $\{z_n\}$, entonces la clausura de esta familia de elementos es un subespacio algebraicamente cerrado.

Tomemos ahora un subespacio cerrado cualquiera A; Schmidt demuestra que si z es un elemento cualquiera del espacio, entonces existen elementos únicos w_1 y w_2 tales que $z = w_1 + w_2$, donde w_1 pertenece a A y w_2 es ortogonal a A, lo que significa que w_2 es ortogonal a todo elemento de A. Este resultado se designa actualmente con el nombre de teorema de la proyección; w_1 es la proyección de z sobre A. Además, $\|w_2\| = \min \|y - z\|$, donde y es un elemento cualquiera de A, y este mínimo se alcanza sólo para $y = w_1$; $\|w_2\|$ recibe el nombre de distancia entre z y A.

Schmidt y Fréchet observaron simultáneamente en 1907 que el espacio de las funciones de cuadrado sumable (con la integral de Lebesgue) tiene una geometría completamente análoga a la del espacio de Hilbert de las sucesiones. Esta analogía quedó aclarada algunos meses después, cuando Riesz, haciendo uso del teorema de Riesz-Fischer (cap. 45, sec. 4) que establece una correspondencia biunívoca entre las funciones medibles de Lebesgue de cuadrado integrable y las sucesiones de números reales de cuadrado sumable, señaló que puede definirse una distancia en el conjunto L^2 de las funciones de cuadrado sumable, y puede usarse esta distancia para construir una geometría sobre este espacio de funciones. Este concepto

de distancia entre dos funciones cualesquiera de cuadrado sumable del espacio L^2 sobre el intervalo $[a,b]$, también fue definido, de hecho, por Fréchet,[18] en la forma

$$\sqrt{\int_a^b [f(x) - g(x)]^2\, dx} \qquad (1)$$

donde la integral está tomada en el sentido de Lebesgue, y considerando que dos funciones que solamente difieren entre sí sobre un conjunto de medida cero, son iguales. El cuadrado de esta distancia recibe también el nombre de desviación cuadrática media de las funciones. El producto escalar de f y g se define de la forma $(f,g) = \int_a^b f(x) \cdot g(x)\, dx$, y las funciones que cumplen $(f,g) = 0$ se llaman ortogonales. La desigualdad de Schwarz

$$\int_a^b f(x)g(x)\, dx \leq \sqrt{\int_a^b f^2\, dx}\ \sqrt{\int_a^b g^2\, dx}$$

y otras propiedades que se verifican en el espacio de las sucesiones de cuadrado sumable, se aplican también al espacio de funciones; en particular, esta clase de funciones de cuadrado sumable constituye un espacio completo. Así pues, se puede identificar el espacio de las funciones de cuadrado sumable y el espacio de las sucesiones de cuadrado sumable que son los coeficientes de Fourier de dichas funciones con respecto a un sistema ortonormal completo de funciones fijo.

Por lo que se refiere a espacios de funciones, tenemos que recordar también (cap. 45, sec. 4) los espacios L^p, $1 < p < \infty$, introducidos por Riesz; estos espacios también son completos con la métrica $d(f_1, f_2) = \left(\int_a^b |f_1 - f_2|^p\, dx\right)^{1/p}$

Aunque volveremos a considerar más adelante el desarrollo de la teoría de espacios abstractos en general, los resultados siguientes se

[18] Comp. Rend, 144, 1907, 1414-1416.

refieren a funcionales y operadores. En su artículo de 1907 ya mencionado, en el que introducía la métrica o «*écart*» para funciones del espacio L^2, y en otro artículo del mismo año,[19] demostró Fréchet que para todo funcional lineal continuo $U(f)$ definido sobre L^2, existe una única función $u(x)$ en L^2, tal que para toda f de L^2,

$$U(f) = \int_a^b f(x)u(x)dx,$$

resultado que generaliza otro obtenido por Hadamard en 1903.[20] En 1909,[21] Riesz generalizó a su vez este resultado, expresando $U(f)$ como una integral de Stieltjes, es decir

$$U(f) = \int_a^b f(x)du(x).$$

Riesz mismo volvió a generalizar este resultado para funcionales lineales A definidos sobre un espacio L^p, que satisfagan la condición de que para todo f en L^p

$$A(f) \leq M\left[\int_a^b |f(x)|^p\, dx\right]^{1/p},$$

donde M depende sólo de A; entonces existe una función $a(x)$ en L^q, única módulo la suma de una función de integral cero, tal que para toda f en L^p

$$U(f) = \int_a^b a(x)f(x)dx. \qquad (2)$$

Este resultado recibe el nombre de teorema de representación de Riesz.

La parte central del análisis funcional se ocupa de la teoría abstracta de los operadores que aparecen en las ecuaciones diferenciales e integrales; este teoría unifica la teoría de autovalores para ecuaciones diferenciales e integrales y para transformaciones lineales que

---

[19] Amer. Math. Soc. Trans., 8, 1907, 433-446.

[20] Comp. Rend., 136, 1903, 351-354 = *Œuvres*, 1, 405-408.

[21] Comp. Rend., 149, 1909, 974-977, y Ann. de l'Ecole Norm. Sup., 28, 1911, 33 ss.

actúan sobre un espacio n-dimensional. Un operador de esta clase, como por ejemplo

$$g(x) = \int_a^b k(x, y)f(y)dy$$

donde k es una función dada, hace corresponder la función g a la f y satisface algunas otras condiciones adicionales. Utilizando la notación A para el operador abstracto, y la notación $g = Af$ para la correspondencia asociada, la linealidad significa que

$$A(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 Af_1 + \lambda_2 Af_2 \qquad (3)$$

donde las λ_i son constantes reales o complejas cualesquiera. La integral indefinida $g(x) = \int_a^x f(t)dt$ y la derivada $f'(x) = Df(x)$ son operadores lineales sobre las correspondientes clases de funciones. La continuidad del operador A significa que si la sucesión de funciones f_n tiende a f en el sentido de la operación límite en el espacio de funciones, entonces Af_n debe tender a Af.

El análogo abstracto de lo que se presenta en el caso de las ecuaciones integrales cuando el núcleo $k(x,y)$ es simétrico, es la propiedad del operador A de ser autoadjunto: si, cualesquiera que sean las funciones f_2 y f_2, se tiene que

$$(Af_1, f_2) = (f_1, Af_2),$$

donde (Af_1,f_2) representa el producto escalar de dos funciones del espacio, entonces A se llama autoadjunto. En el caso de las ecuaciones integrales, si

$$Af = \int_a^b k(x, y)f(y)dy,$$

entonces

$$(Af_1, f_2) = \int_a^b \int_a^b k(x, y)f_1(y)f_2(x)dy\, dx$$

$$(f_1, Af_2) = \int_a^b \int_a^b k(x, y)f_2(y)f_1(x)dy\, dx$$

y $(Af_1, f_2) = (f_1, Af_2)$ si el núcleo es simétrico. Para los operadores autoadjuntos arbitrarios, los autovalores son todos reales, y las autofunciones correspondientes a autovalores distintos son ortogonales entre sí.

Un vigoroso impulso inicial a la teoría abstracta de operadores, que es el núcleo del análisis funcional, se debe a Riesz, en su artículo de 1910 en los *Mathematische Annalen,* donde introduce los espacios L^p (cap. 45, sec. 4). Riesz pretendía generalizar en este artículo la resolución de la ecuación integral

$$\Phi(x) - \lambda \int_a^b K(x, t)\Phi(t)dt = f(x)$$

para funciones en los espacios L^p. Riesz consideró la expresión

$$\int_a^b K(s, t)\Phi(t)\, dt$$

como una transformación actuando sobre la función $\Phi(t)$, la llamó una transformación funcional y la representó por $T(\Phi(t))$. Además, dado que las $\Phi(t)$ que consideraba Riesz estaban en un cierto espacio L^p, la transformación aplicaba funciones en otras del mismo espacio, o de otro; en particular, una transformación u operador que transforme funciones de L^p en funciones de L^p se llama lineal sobre L^p si satisface (3) y si T está acotado, es decir, si existe una constante M tal que para toda f en L^p que satisfaga

$$\int_a^b |f(x)|^p\, dx \leq 1$$

se tiene que

$$\int_a^b |T(f(x))|^p\, dx \leq M^p.$$

Posteriormente, al extremo inferior de tales M se le llamó la norma de T, y se representó por $\|T\|$.

Riesz introdujo también el concepto de operador adjunto o traspuesto de T: para toda g en L^q y todo operador T actuando sobre L^p, se tiene que

$$\int_a^b T(f(x))g(x)\,dx \tag{4}$$

define un funcional sobre L^p, donde g es una función fija y f varía sobre L^p; por tanto, por el teorema de representación de Riesz, existe una función $\psi(x)$ en L^q, única salvo una función cuya integral sea cero, tal que

$$\int_a^b T(f(x))g(x)\,dx = \int_a^b f(x)\psi(x)\,dx \tag{5}$$

El adjunto o traspuesto de T, representado por T^*, se define entonces como aquel operador sobre L^q que, para el T fijo depende sólo de g, y asigna a la función g la ψ de la ecuación (5), es decir, $T^*(g) = \psi$. (En notación moderna, T^* satisface $(Tf,g) = (f,T^*g)$. T^* es una transformación lineal sobre L^q y verifica que $\|T^*\| = \|T\|$.

Riesz consideró entonces la solución de la ecuación

$$T(\Phi(x)) = f(x), \tag{6}$$

donde T es una transformación lineal sobre L^p, la función f es conocida y Φ es desconocida. Demostró que (6) tiene una solución si y sólo si

$$\left|\int_a^b f(x)g(x)dx\right| \leq M\left(\int_a^b |T^*(g(x))|^q\,dx\right)^{1/q}$$

para toda g en L^q. De esta manera se vio conducido al concepto de transformación inversa u operador T^{-1}, y al mismo tiempo para T^{*-1}; con la ayuda del operador adjunto pudo demostrar la existencia de los inversos.

En su artículo de 1910 introduce Riesz la notación

$$\Phi(x) - \lambda K(\Phi(x)) = f(x), \tag{7}$$

donde K representa $\int_a^b K(x,t) * dt$ y donde $*$ debe reemplazarse por la función sobre la que actúa K; sus resultados siguientes se limitan a L^2, donde $K = K^*$. Para tratar el problema de los autovalores de ecuaciones integrales introduce el concepto de continuidad completa

de Hilbert, pero formulado ahora para operadores abstractos. Un operador K sobre L^2 se llama completamente continuo si K transforma toda sucesión de funciones débilmente convergente (cap. 45, sec. 4) en otra fuertemente convergente; es decir, si $\{f_n\}$ converge débilmente, entonces $\{K(f_n)\}$ converge fuertemente. Riesz demostró que el espectro de (7) es discreto (es decir, un K simétrico no tiene espectro continuo) y que las autofunciones asociadas a los autovalores son ortogonales.

Otro enfoque de los espacios abstractos, esta vez usando el concepto básico de norma, fue iniciado unos años más tarde también por Riesz;[22] sin embargo, la definición general de los espacios normados fue dada de una manera casi simultánea durante los años 1920 a 1922 por Stefan Banach (1892-1945), Hans Hahn (1879-1934), Eduard Helly (1884-1943) y Norbert Wiener (1894-1964). Aunque los trabajos de estos matemáticos se solapan en buena parte y resulta difícil decidir en cuestiones de prioridad, el hecho es que la obra de Banach es la que ha tenido mayor influencia. Su motivación fue la generalización de las ecuaciones integrales.

La característica esencial de todos los trabajos que hemos mencionado, y del de Banach en particular,[23] era la de construir un espacio dotado de una norma, pero que en general no era definida a partir de un producto escalar. Mientras que en L^2 se tiene que $\|f\| = (f,f)^{1/2}$, en un espacio de Banach general no se puede definir la norma de esta manera, porque no se tiene previamente un producto escalar.

Banach parte de un espacio E cuyos elementos se representan por x, y, z, ..., mientras que a, b, c, ..., denotan números reales. Los axiomas que debe satisfacer el espacio E se dividen en tres grupos: el primer grupo consta de trece axiomas que expresan el hecho de que E es un grupo abeliano para la suma, que es cerrado para la multiplicación por un escalar real, y que se verifican las propiedades asociativas y distributivas usuales entre las operaciones con números reales y con elementos de E.

El segundo grupo de axiomas caracteriza lo que se llama una norma definida para los elementos (o vectores) de E. La norma es una función con valores reales definida sobre E, que representare-

---

[22] Acta Math., 41, 1918, 71-98.
[23] Fundamenta Mathematicae, 3, 1922, 133-181.

mos por $\|x\|$: para todo número real a y todo vector x de E, la norma tiene las propiedades siguientes:

a) $\|x\| \geqslant 0$.

b) $\|x\| = 0$ si y sólo si $x = 0$.

c) $\|ax\| = |a| \cdot \|x\|$.

d) $\|x + y\| \leqslant \|x\| + \|y\|$.

El tercer grupo consta de un solo axioma, el axioma de completitud, que dice que si $\{x_n\}$ es una sucesión de Cauchy para la norma definida, es decir, si $\lim_{n,p\to\infty} \|x_n - x_p\| = 0$, entonces existe un elemento x en E tal que $\lim_{n\to\infty} \|x_n - x\| = 0$.

Un espacio que satisface estos tres grupos de axiomas recibe el nombre de espacio de Banach o espacio vectorial normado completo. Aunque un espacio de Banach es más general que un espacio de Hilbert, porque no se supone que esté definido en él un producto escalar entre vectores que defina a su vez la norma, como consecuencia de ello se pierde el concepto fundamental de vectores ortogonales en los espacios de Banach que no sean además de Hilbert. Los grupos de condiciones primero y tercero se verifican también en los espacios de Hilbert, pero el segundo grupo es más débil que las condiciones que verifica la norma en un espacio de Hilbert. Los espacios de Banach incluyen a los espacios L^p, los espacios de funciones continuas, los espacios de funciones medibles acotadas, entre otros, siempre que se defina sobre ellos una norma adecuada.

Utilizando el concepto de norma pudo demostrar Banach una serie de resultados conocidos para estos espacios. Uno de los teoremas clave dice lo siguiente: sea $\{x_n\}$ una sucesión de elementos de E tales que

$$\sum_{p=1}^{\infty} \|x_p\| < \infty,$$

entonces $\Sigma_{p=1}^{\infty} x_p$ converge según la norma a un elemento x de E.

Después de demostrar una serie de teoremas, pasó a estudiar Banach los operadores definidos sobre el espacio E pero con valores en otro espacio también de Banach E_1. Un operador F se llama continuo en x_0 con respecto a un conjunto A, si $F(x)$ está definido para todo x de A, x_0 pertenece a A y al conjunto derivado de A, y si $\{x_n\}$ es una sucesión arbitraria de A con límite x_0, entonces $F(x_n)$ converge a $F(x_0)$. También definió Banach la continuidad uniforme

de un operador F con respecto a un conjunto A, para pasar a continuación al estudio de las sucesiones de operadores: la sucesión $\{F_n\}$ de operadores se llama convergente según la norma al operador F sobre un conjunto A, si para todo x de A, $\lim_{n\to\infty} F_n(x) = F(x)$.

Una clase importante de operadores introducida por Banach es la de los operadores aditivos continuos: un operador F se llama aditivo si para todo x y para todo y, $F(x + y) = F(x) + F(y)$. Un operador aditivo continuo verifica la propiedad de que para todo número real a, $F(ax) = aF(x)$; si F es aditivo y continuo en un punto de E, entonces es continuo en todos y además es acotado, es decir, existe una constante M, que depende sólo de F, tal que para todo x en E se tiene que $\|F(x)\| \leqslant M\|x\|$. Otro teorema asegura que si $\{F_n\}$ es una sucesión de operadores aditivos continuos, y si F es un operador aditivo tal que para todo x, $\lim_{n\to\infty} F_n(x) = F(x)$, entonces F es continuo y existe un M tal que para todo n se verifica que $\|F_n(x)\| \leqslant M\|x\|$.

En este artículo demuestra Banach algunos teoremas sobre la resolución de ecuaciones integrales, formulados de una manera abstracta. Si F es un operador continuo que aplica el espacio E en sí mismo, y si existe un número M, con $0 < M < 1$, tal que para todo x' y todo x'' en E, $\|F(x') - F(x'')\| \leqslant M\|x' - x''\|$, entonces existe un único elemento x en E que satisface $F(x) = x$. Aún más importante es el teorema siguiente: consideremos la ecuación

$$x + hF(x) = y \tag{8}$$

donde y es una función conocida de E, F es un operador aditivo continuo de E en E, y h un número real. Sea M el extremo inferior de los números M' que satisfacen $\|F(x)\| \leqslant M' \|x\|$ para todo x; entonces para todo y, y todo valor de h que verifique $|hM| < 1$, existe una función x que satisface la ecuación (8) y además

$$x = y + \sum_{n=1}^{\infty} (-1)^n h^n F^{(n)}(y)$$

donde $F^{(n)}(y) = F(F^{(n-1)}(y))$. Este resultado es una forma del teorema del radio espectral, y una generalización del método de Volterra para resolver ecuaciones integrales.

En 1929 [24] introdujo Banach otro concepto importante del análisis funcional: el concepto de espacio dual o adjunto de un espacio de Banach, idea que también fue introducida de manera independiente por Hahn,[25] pero el trabajo de Banach era más sistemático. Este espacio dual es el espacio de todos los funcionales lineales continuos acotados sobre el espacio dado; la norma para este espacio de funcionales se define como las cotas de los funcionales, con la cual resulta ser un espacio vectorial normado completo, es decir, un espacio de Banach. En realidad, el trabajo de Banach generaliza aquí resultados de Riesz sobre los espacios L^p y L^q, siendo $q = p/p-1$, porque el espacio L^q es equivalente al dual del espacio L^p en el sentido de Banach. La conexión con el trabajo de Banach se pone en evidencia por el teorema de representación de Riesz (2). En otras palabras, el espacio dual de Banach tiene la misma relación con el espacio de Banach dado E, que L^q tiene con L^p.

Banach parte de la definición de un funcional lineal continuo, es decir, una función continua con valores reales definida sobre el espacio E, y demuestra que todo funcional de tal clase está acotado. Un teorema básico, que generaliza otro debido a Hahn, es el que se conoce actualmente como teorema de Hahn-Banach: sea p un funcional con valores reales definido sobre un espacio vectorial normado completo R, que verifica, para todo x e y en R

a) $p(x + y) \leqslant p(x) + p(y)$.

b) $p(\lambda x) = \lambda p(x)$ para $\lambda \geqslant 0$.

Entonces existe un funcional aditivo f sobre R que verifica

$$-p(-x) \leq f(x) \leq p(x)$$

para todo x en R. Hay toda una serie de teoremas sobre el conjunto de los funcionales continuos definidos sobre R.

Las investigaciones sobre funcionales condujeron al concepto de operador adjunto. Sean R y S dos espacios de Banach, y sea U un operador lineal continuo de R en S; sean R^* y S^* los conjuntos de funcionales lineales acotados definidos respectivamente sobre R y S. Entonces U induce una aplicación de S^* en R^*, de la manera siguiente: si g es un elemento de S^*, entonces $g(U(x))$ está bien definido para todo x en R. Por la linealidad de U y de g este es también

[24] Studia Mathematica, 1, 1929, 211-216.

[25] Jour. für Math., 137, 1927, 214-229.

un funcional lineal sobre R, es decir, $g(U(x))$ es un elemento de R^*. Dicho en otras palabras, si $U(x) = y$ entonces $g(y) = f(x)$, donde f es un funcional de R^*. La aplicación inducida U^* de S^* en R^* recibe el nombre de adjunta de la U.

Utilizando este concepto demuestra Banach que si U^* tiene un inverso continuo, entonces $y = U(x)$ tiene solución para todo y en S. También, si $f = U^*(g)$ tiene solución para todo f en R^*, entonces U^{-1} existe y es continuo sobre la imagen de U, siendo la imagen de U en S el conjunto de todos los y para los cuales existe un g en S^* con la propiedad de que $g(y) = 0$ siempre que $U^*(g) = 0$. (Esta última afirmación es una versión generalizada del teorema de la alternativa de Fredholm).

Banach aplicó su teoría de operadores adjuntos a los operadores de Riesz, que había introducido este último en su artículo de 1918. Estos son operadores U de la forma $U = I - \lambda V$ donde I es el operador identidad y V es un operador completamente continuo. Entonces puede aplicarse la teoría abstracta al espacio L^2 de funciones definidas sobre [0,1] y a los operadores

$$U_\lambda(x) = x(s) - \lambda \int_0^1 K(s, t)x(t)dt,$$

donde $\int_0^1 \int_0^1 K^2(s,t)\, ds\, dt < \infty$. La teoría general, aplicada a

$$y(s) = x(s) - \lambda \int_0^1 K(s, t)x(t)dt \tag{9}$$

y a su ecuación traspuesta asociada

$$f(s) = g(s) - \lambda \int_0^1 K(t, s)g(t)dt, \tag{10}$$

nos dice que si λ_0 es un autovalor de (9), entonces λ_0 es también un autovalor de (10) y recíprocamente. Además, (9) tiene un número finito de autofunciones linealmente independientes asociadas a λ_0, y lo mismo ocurre con (10). También ocurre que (9) no tiene solución para todo y cuando $\lambda = \lambda_0$; de hecho, una condición necesaria y suficiente para que (9) tenga una solución, es que se satisfagan las n condiciones

$$\int_0^1 y(t) g_0^{(p)}(t) dt = 0; \qquad p = 1, 2, ..., n,$$

donde $g_0^{(1)}, ..., g_0^{(n)}$ es un conjunto de soluciones linealmente independientes de

$$g(s) - \lambda_0 \int_0^1 K(t, s) g(t) dt = 0.$$

4. La axiomatización de los espacios de Hilbert

La teoría de espacios de funciones y operadores, durante los años veinte, parecía conducir sólo a la abstracción por la abstracción. Incluso Banach no hizo ningún uso concreto de su obra. Este estado de cosas hizo observar más tarde a Hermann Weyl: «No fue mérito alguno, sino favor de la fortuna el que se descubriese, a partir de 1923, que la teoría espectral del espacio de Hilbert era el instrumento matemático adecuado a la mecánica cuántica». La investigación en mecánica cuántica mostró que los observables de un sistema físico se pueden representar por operadores lineales simétricos en un espacio de Hilbert, y que los autovalores y autovectores (autofunciones) del operador particular que representa la energía son los niveles de energía de un electrón en el átomo y los correspondientes estados cuánticos estacionarios del sistema. Las diferencias entre dos autovalores dan las frecuencias de los cuantos de luz emitidos, y definen así el espectro de radiación de la sustancia. En 1926 Erwin Schrödinger presentó su teoría cuántica basada en ecuaciones diferenciales, y demostró la identidad de esta última teoría con la teoría de matrices infinitas de Werner Heisenberg (1925), que éste había aplicado a la teoría cuántica. Sin embargo, faltaba una teoría general que unificase los trabajos de Hilbert y la teoría de autofunciones para ecuaciones diferenciales.

El uso de operadores en la teoría cuántica estimuló el trabajo en una teoría abstracta del espacio de Hilbert y sus operadores, tarea que fue emprendida en primer lugar por John von Neumann (1903-1957) en 1927. Su planteamiento incluye a la vez el espacio de las sucesiones de cuadrado sumable y el espacio L^2 de funciones definidas en un intervalo común.

En dos largos artículos [26] presentó von Neumann un tratamiento axiomático del espacio de Hilbert y de los operadores sobre estos espacios. Aunque puede detectarse el origen de los axiomas de von Neumann en los trabajos de Norbert Wiener, Weyl y Banach, la obra de von Neumann era más completa y ejerció una mayor influencia; su objetivo principal era formular una teoría general de autovalores para una amplia gama de operadores llamados hermíticos.

Von Neumann introduce un espacio L^2 de funciones complejas, medibles y de cuadrado integrable, definidas sobre un conjunto medible arbitrario E del plano complejo, así como el espacio análogo o espacio de sucesiones complejas, es decir, el conjunto de todas las sucesiones de números complejos a_1, a_2, ... con la propiedad de que $\sum_{p=1}^{\infty}|a_p|^2 < \infty$. El teorema de Riesz-Fischer permitía demostrar entonces que existe una correspondencia biunívoca entre las funciones del primer espacio y las sucesiones del segundo espacio, definida de la forma siguiente. Elijamos en el espacio de funciones una sucesión ortonormal completa de funciones $\{\Phi_n\}$; entonces si f es una función cualquiera del espacio, los coeficientes de Fourier de f en su desarrollo con respecto a las $\{\Phi_n\}$ forman una sucesión del espacio de sucesiones, y recíprocamente, si uno parte de una sucesión así, entonces hay una única función de L^2 (salvo una función cuya integral sea cero) que tiene esta sucesión como la de sus coeficientes de Fourier con respecto a las $\{\Phi_n\}$.

Además, si se define el producto escalar (f,g) en el espacio de funciones por

$$(f, g) = \int_E f(z)\,\overline{g(z)}dz$$

donde $\overline{g(z)}$ es el complejo conjugado de $g(z)$, y en el espacio de sucesiones el producto escalar de dos sucesiones a y b por

$$(a, b) = \sum_{p=1}^{\infty} a_p\bar{b}_p,$$

entonces, si f corresponde a a y g a b, se tiene $(f,g) = (a,b)$.

Un operador hermítico R sobre uno de estos espacios se define como un operador lineal con la propiedad de que, para todas las f

---

[26] Math. Ann., 102, 1929-1930, 49-131 y 370-427 = *Coll. Works*,2, 3-143.

y g pertenecientes a su dominio, $(Rf,g) = (f,Rg)$; análogamente, en el espacio de las sucesiones debe verificarse $(Ra,b) = (a,Rb)$.

La teoría de Von Neumann ofrece un tratamiento axiomático para ambos espacios a la vez, el de funciones y el de sucesiones; los axiomas que propone son los siguientes:

A) H es un espacio vectorial; es decir, están definidas sobre H un suma y una multiplicación por escalares, de manera que si f_1, f_2 son elementos de H y a_1 y a_2 son números complejos cualesquiera, entonces $a_1f_1 + a_2f_2$ es otro elemento de H.

B) Está definido sobre H un producto escalar, o función definida para todas las parejas de vectores f y g y con valores complejos, representada por (f,g), con las propiedades:

a) $(af,g) = a(f,g)$.
b) $(f_1 + f_2,g) = (f_1,g) + (f_2,g)$.
c) $(f,g) = \overline{(g,f)}$.
d) $(f,f) \geqslant 0$.
e) $(f,f) = 0$ si y sólo si $f = 0$.

Dos elementos f y g se llaman ortogonales si $(f,g) = 0$; la norma de f, representada por $\|f\|$, se define como $\sqrt{(f,f)}$, y $\|f - g\|$ define una métrica sobre el espacio H.

C) Con la métrica que acabamos de definir, H es separable. Es decir, existe en H un subconjunto denso numerable con respecto a la métrica $\|f - g\|$.

D) Para todo entero positivo n, hay en H n vectores linealmente independientes.

E) H es completo. Es decir, si la sucesión $\{f_n\}$ es tal que $\|f_n - f_m\|$ tiende a cero cuando m y n tienden a ∞, entonces hay un f en H tal que $\|f - f_n\|$ tiende a cero cuando n tiende a ∞. (Esta convergencia es equivalente a la convergencia fuerte.)

De estos axiomas se deducen inmediatamente algunas propiedades sencillas, como son la desigualdad de Schwarz $\|(f,g)\| \leqslant \|f\| \cdot \|g\|$, el hecho de que todo conjunto ortonormal completo de elementos de H tiene que ser numerable, la desigualdad de Parseval, es decir, $\Sigma_{p=1}^{\infty} \|(f,\phi_p)\|^2 \geqslant \|f\|^2$, etc.

A continuación von Neumann pasa a estudiar los subespacios vectoriales de H y los operadores proyección. Si M y N son subespacios cerrados de H, entonces se define $M - N$ como el conjunto de todos los elementos de M que son ortogonales a todos los elementos de N. Entonces, el teorema de la proyección dice: sea M un subespacio cerrado de H; entonces cualquier vector f de H puede

expresarse de una y sólo una manera como $f = g + h$, donde g está en M y h en $H - M$. El operador proyección asociado a M, P_M, se define de la forma $P_M(f) = g$; es decir, es el operador, definido sobre todo H, que proyecta cada elemento f en su componente según M.

En el segundo de los artículos citados introduce von Neumann dos topologías sobre H, la fuerte y la débil. La topología fuerte es simplemente la topología métrica definida a partir de la norma. La topología débil, en la que no entraremos aquí, da el sistema de entornos asociados a la convergencia débil.

Von Neumann presenta un buen número de resultados sobre operadores en el espacio de Hilbert. Se presupone que todos los operadores estudiados son lineales, y la integración se entiende generalmente en el sentido de Lebesgue-Stieltjes. Una transformación u operador lineal acotado es uno que transforma elementos de un espacio de Hilbert en otro, satisface la condición de linealidad (3), y es acotado, es decir, existe un número M tal que para todo f del espacio sobre el que actúa el operador R,

$$\|R(f)\| \leq M\|f\|.$$

El ínfimo de los M posibles recibe el nombre de módulo (norma) de R. Esta última condición es equivalente a la continuidad del operador. También es fácil ver que la continuidad en un único punto, junto con la linealidad, es suficiente para garantizar la continuidad en todos los puntos, y por tanto la acotación.

Está también el operador adjunto de R, R^*. Para los operadores hermíticos $R = R^*$, es decir, R es autoadjunto. Si $RR^* = R^*R$ entonces se dice que R es normal. Si $RR^* = R^*R = I$, donde I es el operador identidad, entonces R es análogo a una transformación ortogonal, y se llama unitario; para todo R unitario, $\|F(f)\| = \|f\|$.

Otro de los resultados de von Neumann afirma que si R es hermítico sobre el espacio de Hilbert y satisface la condición de cierre débil, a saber, que si f_n tiende a f y $R(f_n)$ tiende a g entonces $R(f) = g$, y entonces R es acotado. Además, si R es hermítico, entonces el operador $I - \lambda R$ tiene inverso para todo λ real o complejo exterior a un intervalo (m,M) del eje real en el que están los valores de $(R(f),f)$ cuando $\|f\| = 1$. Otro resultado más importante es el de que a todo operador lineal acotado y hermítico le corresponden otros dos operadores E_- y E_+ *(Einzeltransformationen)* con las siguientes propiedades:

a) $E_- E_- = E_-$, $E_+ E_+ = E_+$, $I = E_- + E_+$.

b) E_- y E_+ conmutan, y también conmutan con todo operador que conmute con R.

c) RE_- y RE_+ son respectivamente negativo y positivo (RE_+ es positivo si $(RE_+(f),f) \geqslant 0$).

d) Para todo f tal que $R(f) = 0$, $E_-(f) = 0$ y $E_+(f) = f$.

Von Neumann estableció también una importante relación entre los operadores hermíticos y unitarios, la de que si U es unitario y R es hermítico, entonces $U = e^{iR}$. Más adelante también generalizó su teoría a operadores no acotados. Pese a que sus contribuciones en este campo, como las de otros matemáticos, fueron de la mayor importancia, una exposición detallada de ellas nos llevaría demasiado lejos y nos introduciría en desarrollos recientes más complejos.

Se han hecho, y se siguen haciendo, aplicaciones del análisis funcional al problema generalizado de los momentos, a la mecánica estadística, a los teoremas de existencia y unicidad en ecuaciones en derivadas parciales, a los teoremas del punto fijo, y a muchos otros. Hoy día el análisis funcional juega un papel importante en el cálculo de variaciones y en la teoría de representaciones de grupos continuos compactos, así como en álgebra, cálculo aproximado, topología y teoría de funciones de variable real. A pesar de esta diversidad de aplicaciones, hay una deplorable ausencia de aplicaciones nuevas a los grandes problemas del análisis clásico. Este fracaso desilusionó a los fundadores del análisis funcional.

Bibliografía

Bernkopf, M.: «The Development of Function Spaces with Particular Reference to their Origins in Integral Equation Theory». *Archive for History of Exact Sciences*, 3, 1966, 1-96.

— «A History of Infinite Matrices». *Archive for History of Exact Sciences*, 4, 1968, 308-358.

Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.

Dresden, Arnold: «Some Recent Work in the Calculus of Variations». *Amer. Math. Soc. Bull.*, 32, 1926, 475-521.

Fréchet, M.: *Notice sur les travaux scientifiques de M. Maurice Fréchet*, Hermann, 1933.

Hellinger, E., y Toeplitz: «Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, vol. 2, part. III, 2.ª half, 1335-1597.
Hildebrandt, T. H.: «Linear Functional Transformations in General Spaces». *Amer. Math. Soc. Bull.*, 37, 1931, 185-212.
Lévy, Paul: «Jacques Hadamard, sa vie et son Œuvre». *L'Enseignement Mathématique* (2), 13, 1967, 1-24.
McShane, E. J.: «Recent Developments in the Calculus of Variations». Amer. Math. Soc. Semicentennial Publications, 2, 1938, 69-97.
Neumann, John von: *Collected Works*, Pergamon Press, 1961, vol. 2.
Sanger, Ralph G.: «Functions of Lines and the Calculus of Variations». *University of Chicago Contributions to the Calculus of Variations for 1931-1932*, University of Chicago Press, 1933, 193-293.
Tonelli, L.: «The Calculus of Variations». *Amer. Math. Soc. Bull.*, 31, 1925, 163-172.
Volterra, Vito: *Opere matematiche*, 5 vols., Accademia Nazionale dei Lincei, 1954-1962.

Capítulo 47
LA TEORIA DE SERIES DIVERGENTES

> Es realmente una extraña circunstancia de nuestra ciencia el que aquellas series que a principios de siglo se suponía que habían de ser desterradas de una vez y para siempre de la matemática rigurosa aparezcan, a su final, llamando a la puerta para ser readmitidas.
>
> JAMES PIERPONT

> Esta serie es divergente, por tanto algo podremos hacer con ella.
>
> OLIVER HEAVISIDE

1. Introducción

La toma en consideración desde finales del siglo XIX en adelante de un tema como el de las series divergentes indica cuán radicalmente los matemáticos han revisado su propia consideración de la naturaleza de la matemática. Mientras que en la primera parte del siglo XIX aceptaron la exclusión de las series divergentes sobre la base de que la matemática se encontraba restringida por alguna necesidad intrínseca o por los dictados de la naturaleza a una clase fija de conceptos correctos, a finales del siglo reconocían su libertad para aceptar cualquier idea que pareciera presentar alguna utilidad.

Cabe recordar que las series divergentes se utilizaron a lo largo de todo el siglo XVIII, con un reconocimiento más o menos consciente de su divergencia, porque suministraban aproximaciones útiles de funciones tomando unos pocos términos solamente. A partir de la aparición de la matemática rigurosa con Cauchy, la mayor parte de los matemáticos siguieron sus criterios y rechazaron las series divergentes por desconfiar de ellas. Sin embargo, unos pocos matemáticos (cap. 40, sec. 7) continuaron defendiendo las series divergentes porque estaban convencidos de su utilidad, bien para el cálculo

de funciones o como representación analítica de las funciones de las que se derivaban. Otros las defendían como método de descubrimiento. Así, por ejemplo, De Morgan [1] decía que «Tenemos que admitir que muchas series son tales que no podemos utilizarlas de momento con seguridad, excepto como método de descubrimiento, cuyos resultados tendrán que ser comprobados posteriormente, y sin duda incluso el enemigo más acérrimo de las series divergentes hace este uso de ellas en privado».

Los astrónomos siguieron utilizando series divergentes, después incluso de que fueran excluidas, por las necesidades de su ciencia a efectos de cálculo. Dado que unos pocos de los primeros términos de tales series daban una aproximación numérica útil, los astrónomos solían ignorar el hecho de que la serie completa era divergente, mientras que los matemáticos, interesados en el comportamiento no de los primeros diez o veinte términos, sino de la serie total, no podían aceptarlas únicamente a causa de su utilidad.

Sin embargo, como ya hemos hecho observar (cap. 40, sec. 7) tanto Abel como Cauchy no dejaron de ser conscientes de que, al excluir las series divergentes, estaban eliminando algo útil. Cauchy no sólo continuó usándolas (véase más abajo) sino que escribió incluso un artículo con el título «Sur l'emploi légitime des séries divergentes»,[2] en el que, hablando de la serie de Stirling para log $\Gamma(x)$ o log $m!$ (cap. 20, sec. 4), señala que la serie, aun siendo divergente para todos los valores de x, puede ser utilizada para calcular log $\Gamma(x)$ para x positivo y suficientemente grande. De hecho, Cauchy demuestra que, fijado el número n de términos tomados, el error absoluto cometido al sumarlos es menos que el valor absoluto del siguiente término, y el error se va haciendo más pequeño según x aumenta. Cauchy trató de explicarse por qué era tan buena la aproximación dada por la serie, pero no tuvo éxito.

La utilidad de las series divergentes terminó por convencer a los matemáticos de que debían tener alguna característica que, una vez aislada, mostraría por qué daban tan buenas aproximaciones. Como decía Oliver Heaviside en el segundo volumen de su *Electromagnetic Theory* (1899), «Debo decir unas pocas palabras sobre la diferenciación generalizada y las series divergentes... No es fácil levantar entusiasmo alguno una vez que ha sido enfriado artificialmente por los

---

[1] Trans Cambridge Phil. Soc., 8, Part II, 1844, 182-203, pub. 1849.

[2] Comp. Rend., 17, 1843, 370-376 = *Œuvres*, (1), 8, 18-25.

aguafiestas de los rigoristas... Tendrá que haber una teoría de series divergentes, o digamos una teoría más general de funciones que la actual, que incluya las series convergentes y divergentes en una totalidad armoniosa». Heaviside no sabía, cuando hizo esta observación, que ya se habían dado los primeros pasos para ello.

La buena disposición de los matemáticos para dedicarse al estudio de las series divergentes se vio sin duda reforzada por otra circunstancia que había ido penetrando de manera gradual en la atmósfera matemática de la época: la geometría no euclídea y las nuevas álgebras. Los matemáticos comenzaron a darse cuenta lentamente de que la matemática es obra del hombre y que la definición de convergencia de Cauchy ya no podía considerarse como una necesidad impuesta por algún poder sobrehumano. A finales del siglo XIX consiguieron aislar la propiedad esencial de aquellas series divergentes que dan aproximaciones útiles a funciones. A dichas series las llamó Poincaré asintóticas, aunque durante todo el siglo recibieron el nombre de semiconvergentes, término introducido por Legendre en su *Essai sur la théorie des nombres* (1798, p. 13), y utilizado también para las series oscilantes.

La teoría de series divergentes se ocupa de dos cuestiones importantes. La primera es la que ya hemos mencionado brevemente, a saber, que algunas de estas series pueden, con un número fijo de términos aproximar una función cada vez mejor según aumenta la variable. De hecho Legendre, en su *Traité des fonctions elliptiques* (1825-1828), ya había caracterizado tales series por la propiedad de que el error cometido al sumar hasta un término cualquiera es del orden del primer término omitido. La segunda cuestión de la teoría de las series divergentes es el concepto de sumabilidad. Es posible definir la suma de una serie de manera completamente nueva y que dé sumas finitas para series divergentes en el sentido de Cauchy.

2. La utilización informal de las series divergentes

Ya hemos tenido ocasión de hablar de la obra de matemáticos del siglo XVIII que utilizaron tanto series convergentes como divergentes. Durante el siglo XIX, antes y después de que Cauchy prohibiera las series divergentes, hubo matemáticos y físicos que siguieron usándolas. Una aplicación nueva consistió en el cálculo de integrales por desarrollo en serie. Naturalmente los matemáticos no eran cons-

cientes en esa época de que lo que estaban calculando eran desarrollos en series asintóticas completas o los primeros términos de desarrollos de integrales en series asintóticas.

El cálculo asintótico de integrales se remonta a Laplace por lo menos. En su *Théorie analytique des probabilités* (1812)[3] Laplace obtiene integrando por partes el desarrollo de la función error

$$Erfc(T)=\int_T^\infty e^{-t^2}dt=\frac{e^{-T^2}}{2T}\left\{1-\frac{1}{2T^2}+\frac{1\cdot 3}{(2T^2)^2}-\frac{1\cdot3\cdot5}{(2T^2)^3}+\ldots\right\}.$$

Hace notar que la serie es divergente, pero la utiliza para calcular Erfc (T) para valores grandes de T.

En el mismo libro[4] también señala Laplace que

$$\int \phi(x)\,\{u(\,x)\}^2\,dx$$

para valores de s grandes depende de los valores de $u(x)$ próximos a sus puntos estacionarios, es decir, de los valores de x para los que $u'(x) = 0$. Laplace utiliza esta observación para demostrar que

$$s! \sim s^{s+1/2}e^{-s}\sqrt{2\pi}\left(1+\frac{1}{12s}+\frac{1}{288s^2}+\ldots\right),$$

resultado que también puede obtenerse a partir de la aproximación de Stirling de log $s!$ (cap. 20, sec. 4).

En su *Théorie analytique des probabilités,*[5] tuvo ocasión Laplace de estudiar integrales de la forma

$$f(x) = \int_a^b g(t)e^{xh(t)}\,dt \tag{1}$$

donde g puede ser una función compleja, h y t son reales, y x es positivo y grande. Laplace observó que la contribución principal a la integral proviene del entorno inmediato de aquellos puntos del dominio de integración donde $h(t)$ alcanza su máximo absoluto. Esta

---

[3] Tercera edición, 1820, 88-109 = *Œuvres*, 7, 89-110.

[4] *Œuvres*, 7, 128-131.

[5] Tercera edición, 1820, vol. 1, parte 2, Cap. 1 = *Œuvres*, 7, 89-110.

contribución de la que habla Laplace es el primer término de lo que hoy llamaríamos un desarrollo en serie asintótica de la integral. Si $h(t)$ tiene exactamente un máximo en $t = a$, entonces el resultado de Laplace se reduce a

$$f(x) \sim g(a)e^{xh(a)} \sqrt{\frac{-\pi}{2xh''(a)}}$$

cuando x tiende a ∞.

Si en lugar de (1) la integral que hay que calcular es

$$f(x) = \int_a^b g(t)e^{ixh(t)}\, dt \tag{2}$$

para t y x reales y x grande, entonces $|e^{i\,xh(t)}|$ es constante y no se puede aplicar el método de Laplace. En este caso es aplicable, sin embargo, un método vislumbrado por Cauchy en el más importante de sus artículos sobre la propagación de las ondas,[6] hoy llamado principio de la fase estacionaria. Este principio afirma que la contribución más importante a la integral proviene del entorno inmediato de los puntos estacionarios de $h(t)$, es decir, de aquellos en que $h'(t) = 0$. Este principio resulta intuitivamente razonable porque podemos considerar el integrando como una corriente oscilante u onda de amplitud $|g(t)|$; si t es el tiempo, entonces la velocidad de la onda es proporcional a $xh'(t)$, y si $h'(t) \neq 0$ esta velocidad crece indefinidamente cuando x tiende a infinito. Las oscilaciones son tan rápidas entonces que durante un periodo completo $g(t)$ es aproximadamente constante y $xh(t)$ es aproximadamente lineal, de manera que la integral sobre un período completo se anula. Este razonamiento falla para los valores de t tales que $h'(t) = 0$. Así pues, lo más verosímil es que los puntos estacionarios de $h(t)$ den la contribución principal al valor asintótico de $f(x)$. Si τ es un valor de t en el que $h'(t) = 0$ y $h''(\tau) < 0$, entonces

$$f(x) \sim \sqrt{\frac{2\pi}{xh''(\tau)}}\, g(\tau)e^{ixh(\tau)+i\pi/4}$$

cuando x tiende a infinito.

[6] Mém. de l'Acad. des Sci., Inst. France, 1, 1827, Note 16 = *Œuvres*, (1), 1, 230.

Este principio fue utilizado por Stokes para calcular la integral de Airy (véase más adelante) en un artículo de 1856 [7] y lo formuló explícitamente Lord Kelvin.[8] Sin embargo, la primera demostración satisfactoria de dicho principio se debe a George N. Watson (1886-1965).[9]

Durante las primeras décadas del siglo XIX, tanto Cauchy como Poisson calcularon muchas integrales dependientes de un parámetro por desarrollo en serie de potencias del parámetro. En el caso de Poisson las integrales aparecían en problemas geofísicos de conducción del calor y de transmisión de vibraciones elásticas, mientras que Cauchy estudiaba problemas de ondas acuáticas, óptica y astronomía. Por ejemplo, trabajando sobre la difracción de la luz [10] da Cauchy expresiones en series divergentes de las integrales de Fresnel

$$\int_0^m \cos\left(\frac{\pi}{2}z^2\right)dz = \frac{1}{2} - N\cos\frac{\pi}{2}m^2 + M\operatorname{sen}\frac{\pi}{2}m^2$$

$$\int_0^m \operatorname{sen}\left(\frac{\pi}{2}z^2\right)dz = \frac{1}{2} - M\cos\frac{\pi}{2}m^2 - N\operatorname{sen}\frac{\pi}{2}m^2,$$

donde

$$M = \frac{1}{m\pi} - \frac{1\cdot 3}{m^5\pi^3} + \frac{1\cdot 3\cdot 5\cdot 7}{m^9\pi^5} - \dots, \qquad N = \frac{1}{m^3\pi^2} - \frac{1\cdot 3\cdot 5}{m^7\pi^4} +$$

A lo largo del siglo XIX se inventaron otros diversos métodos para calcular integrales, tal como el de descenso más rápido. La teoría completa de todos estos métodos y la comprensión correcta de a qué conducía la aproximación, fueran términos aislados o series completas, tuvo que esperar a la creación de la teoría de series asintóticas.

Muchas de las integrales que se desarrollaron por los métodos anteriores aparecieron por primera vez como soluciones de ecuacio-

---

[7] Trans. Cambridge Phil. Soc., 9, 1856, 166-187 = *Math. and Phys. Papers*, 2, 329-357.

[8] Phil. Magaz., (5), 23, 1887, 252-255 = *Math. and Phys. Papers*, 4, 303-306.

[9] Proc. Camb. Phil. Soc., 19, 1918, 49-55.

[10] Comp. Rend., 15, 1842, 554-556 y 573-578 = *Œuvres*, (1), 7, 149-157.

nes diferenciales. Otro uso de las series divergentes fue para resolver directamente ecuaciones diferenciales, uso que puede hacerse remontar al menos a la obra de Euler [11] donde, al tratar de resolver el problema de la cuerda vibrante no uniforme (cap. 22, párr. 3) da la solución en serie asintótica de una ecuación diferencial ordinaria que esencialmente se reduce a la ecuación de Bessel de orden $r/2$, para r entero.

Jacobi [12] dio el desarrollo asintótico de $J_n(x)$ para x grande:

$$J_n(x) \sim \left(\frac{2}{\pi x}\right)^{1/2}\left[\cos\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right)\left\{1 - \frac{(4n^2-1^2)(4n^2-3^2)}{2!\,(8x)^2} + \dots\right\} - \operatorname{sen}\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right)\left\{\frac{4n^2-1^2}{1!\,8x} - \frac{(4n^2-1^2)(4n^2-3^2)(4n^2-5^2)}{3!\,(8x)^3} + \dots\right\}\right].$$

Liouville introdujo una aplicación algo distinta de las series divergentes a la resolución de ecuaciones diferenciales. Buscaba soluciones aproximadas [13] de la ecuación diferencial

$$\frac{d}{dx}\left(p\,\frac{dy}{dx}\right) + (\lambda^2 q_0 + q_1)y = 0, \qquad (3)$$

donde p, q_0 y q_1 son funciones positivas de x, λ un parámetro y se busca la solución para $a \leqslant x \leqslant b$. Aquí, en contra de lo que ocurría en los problemas de contorno donde se buscaban valores discretos de λ, Liouville estaba interesado en conseguir alguna forma aproximada de y para valores grandes de λ. Para ello introduce las variables

$$t = \int_{x_0}^{x}\left(\frac{q_0}{p}\right)^{1/2} dx, \qquad w = (q_0 p)^{1/4} y \qquad (4)$$

obteniendo

$$\frac{d^2 w}{dt^2} + \lambda^2 w = rw, \qquad (5)$$

[11] Novi Comm. Acad. Sci. Petrop., 9, 1762-1763, 246-304, pub. 1764 = *Opera*, (2), 10, 293-343.

[12] Astronom. Nach., 28, 1849, 65-94 = *Werke*, 7, 145-174.

[13] Jour. de Math., 2, 1837, 16-35.

donde

$$r = (q_0p)^{-1/4} \frac{d^2}{dt^2} (q_0p)^{1/4} - \frac{q_1}{q_0}.$$

y a continuación utiliza un proceso que se reduce, en términos modernos, a la resolución por aproximaciones sucesivas de una ecuación integral del tipo de Volterra, que es la siguiente:

$$w(t) = c_1 \cos \lambda t + c_2 \operatorname{sen} \lambda t + \int_{t_0}^{t} \frac{\operatorname{sen} \lambda(t-s)}{\lambda} r(s)w(s)\, ds.$$

Liouville afirma ahora que, para valores suficientemente grandes de λ, la primera aproximación a las soluciones de (5) sería

$$w \sim c_1 \cos \lambda t + c_2 \operatorname{sen} \lambda t. \tag{6}$$

Utilizando ahora los valores de w y de t dados por (4) para obtener la solución aproximada de (3), resulta de (6) que

$$y \sim c_1 \frac{1}{(q_0p)^{1/4}} \cos\left\{\lambda \int_{x_0}^{x} \left(\frac{q_0}{p}\right)^{1/2} dx\right\} + c_2 \frac{1}{(q_0p)^{1/4}} \operatorname{sen}\left\{\lambda \int_{x_0}^{x} \left(\frac{q_0}{p}\right)^{1/2} dx\right\}. \tag{7}$$

Aunque Liouville no lo sabía, lo que había obtenido era el primer término de una solución de (3) en serie asintótica, para valores grandes de λ.

Green [14] utilizó el mismo método en el estudio de la propagación de ondas en un canal. Este método se ha visto ligeramente generalizado para ecuaciones de la forma

$$y'' + \lambda^2 q(x, \lambda) y = 0 \tag{8}$$

donde λ es un parámetro positivo grande y x puede ser real o complejo. Las soluciones suelen expresarse como

$$y \sim q^{-1/4} \exp\left(\pm i\lambda \int_{0}^{x} q^{1/2}\, dx\right)\left[1 + O\left(\frac{1}{\lambda}\right)\right]. \tag{9}$$

---

[14] Trans. Camb. Phil. Soc., 6, 1837, 457-462 = *Math. Papers*, 225-230.

El término de error $O(1/\lambda)$ implica que la solución exacta incluiría un término $F(x,\lambda)/\lambda$ donde $|F(x,\lambda)|$ está acotado para todo x del dominio considerado y para $\lambda > \lambda_0$. La forma del término de error es válida para un dominio restringido del plano complejo x. Ni Liouville ni Green dieron el término de error o las condiciones bajo las que sus soluciones eran válidas. La aproximación más general y precisa (9) aparece explícitamente en artículos de Gregor Wentzel (1898)[15], Hendrik A. Kramers (1894-1952),[16] León Brillouin (1889-1969) [17] y Harold Jeffreys (1891),[18] y se la conoce familiarmente como la solución *WKBJ*. Todos ellos fueron físicos que trabajaron en teoría cuántica con la ecuación de Schrödinger.

En un artículo leído en 1850 [19] estudia Stokes el valor de la integral de Airy

$$W = \int_0^\infty \cos \pi/2\ (w^3 - mw)\ dw \qquad (10)$$

para $|m|$ grande. Esta integral representa la intensidad de la luz difractada en las proximidades de una cáustica. Airy había dado una serie para W en potencias de m que, a pesar de ser convergente para todo m, no resultaba útil para el cálculo de $|m|$ grande. El método de Stokes consistía en construir una ecuación diferencial de la que la integral es una solución particular, y después resolver dicha ecuación diferencial en términos de series divergentes que pudieran ser útiles para el cálculo (a estas series las llamó semiconvergentes).

Una vez demostrado que $U = (\pi/2)^{1/3}\ W$ satisface la ecuación diferencial de Airy

$$\frac{d^2U}{dn^2} + \frac{n}{3}U = 0, \qquad n = \left(\frac{\pi}{2}\right)^{2/3} m, \qquad (11)$$

demuestra Stokes que para n positivo

[15] Zeit. fur Physik, 38, 1926, 518-529.
[16] Zeit. fur Physik, 39, 1926, 828-840.
[17] Comp. Rend., 183, 1926, 24-26.
[18] Proc. London Math. Soc., (2), 23, 1923, 428-436.
[19] Trans. Camb. Phil. Soc., 9, 1856, 166-187 = *Math. and Phys. Papers*, 2, 329-357.

$$U = An^{-1/4}\left(R\cos\frac{2}{3}\sqrt{\frac{n^3}{3}} + S\,\text{sen}\,\frac{2}{3}\sqrt{\frac{n^3}{3}}\right) + Bn^{-1/4}\left(R\,\text{sen}\,\frac{2}{3}\sqrt{\frac{n^3}{3}} - S\cos\frac{2}{3}\sqrt{\frac{n^3}{3}}\right), \tag{12}$$

donde

$$R = 1 - \frac{1\cdot5\cdot7\cdot11}{1\cdot2\cdot16^2\cdot3n^3} + \frac{1\cdot5\cdot7\cdot11\cdot13\cdot17\cdot19\cdot23}{1\cdot2\cdot3\cdot4\cdot16^4\cdot3^2n^6}\ \ldots, \tag{13}$$

$$S = \frac{1\cdot5}{1\cdot16(3n)^{1/2}} - \frac{1\cdot5\cdot7\cdot11\cdot13\cdot17}{1\cdot2\cdot3\cdot16^3(3n^3)^{3/2}} + \ldots \tag{14}$$

Los valores de A y B que hacen que U dé la integral vienen determinados por un argumento especial (donde Stokes utiliza el principio de la fase estacionaria) y da también un resultado análogo para n negativo.

La serie (12) y la que corresponde a n negativo se comportan como series convergentes para cierto número de términos, pero de hecho divergen, y Stokes observó que podían ser utilizadas para el cálculo aproximado. Dado un valor de n, pueden utilizarse los términos desde el primero hasta uno que resulte más pequeño para dicho valor de n. Stokes dio un razonamiento de tipo cualitativo para mostrar por qué las series son útiles para el cálculo numérico.

Stokes mismo se encontró con una dificultad especial en la resolución de (11) para n positivo y negativo. No pudo pasar de la serie para n positivo a la serie para n negativo simplemente haciendo variar n a través de cero porque dicha serie no tenía sentido para $n = 0$. En vista de eso intentó pasar de valores de n positivos o negativos a través de valores complejos, pero esto no da la serie correcta y los multiplicadores constantes.

Lo que descubrió Stokes, no sin esfuerzo,[20] fue que si para un cierto dominio de la amplitud de n, una solución general era representada por cierta combinación lineal de dos series asintóticas, cada una de las cuales es una solución, entonces no necesariamente la misma combinación lineal de los dos desarrollos asintóticos funda-

[20] Trans. Camb. Phil. Soc., 10, 1857, 106-128 = *Math. and Phys. Papers*, 4, 77-109.

mentales representa la misma solución general en un entorno de ese dominio de amplitudes. Stokes vio que las constantes de la combinación lineal cambian abruptamente cuando se atraviesa las curvas n = const., ahora llamadas curvas de Stokes.

Aunque Stokes se había interesado en principio en el cálculo de integrales, vio claramente que se podrían utilizar las series divergentes para resolver ecuaciones diferenciales en general. Mientras que Euler, Poisson y otros habían resuelto ecuaciones concretas de esta manera, sus resultados parecían reducirse a trucos que daban respuesta a problemas físicos concretos. Stokes dio varios ejemplos en sus artículos de 1856 y 1857.

Los trabajos anteriores sobre el cálculo de integrales y la resolución de ecuaciones diferenciales por medio de series divergentes constituyen una buena muestra de la obra de muchos matemáticos y físicos.

3. La teoría formal de las series asintóticas

Poincaré y Stieltjes lograron en 1886, de manera independiente, una definición formal y una caracterización completa de aquellas series divergentes que resultan útiles para la representación y cálculo de funciones. Poincaré llamó a estas series asintóticas, mientras que Stieltjes continuó utilizando el nombre de semiconvergentes. Poincaré se ocupó del tema motivado por sus investigaciones sobre la resolución de ecuaciones diferenciales lineales. Impresionado por la utilidad de las series divergentes en astronomía, trató de determinar cuáles eran útiles y por qué, logrando al fin aislar y formular la propiedad esencial. Una serie de la forma

$$a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots, \tag{15}$$

donde los a_i son independientes de x, se dice que representa asintóticamente la función $f(x)$ para valores grandes de x, siempre que

$$\lim_{x \to \infty} x^n \left[f(x) - \left(a_0 + \frac{a_1}{x} + \ldots + \frac{a_n}{x^n} \right) \right] = 0 \tag{16}$$

[21] Acta Math., 8, 1886, 293-344 = *Œuvres*, 1, 290-332.

para $n = 0, 1, 2, 3, \ldots$. La serie (15) será en general divergente, pero en casos especiales puede converger. La relación que acabamos de definir entre la función $f(x)$ y la serie se representa por

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots .$$

Tales series vienen a ser desarrollos de funciones en un entorno de $x = \infty$. En su artículo de 1886 Poincaré se limitó a valores de x reales; sin embargo, la definición sirve también para x compleja si sustituimos $x \to \infty$ por $|x| \to \infty$, aunque la validez de la representación puede verse ahora limitada a un sector del plano complejo con vértice en el origen.

Como hemos dicho, la serie (15) es asintótica a $f(x)$ en un entorno de $x = \infty$. Sin embargo, la definición ha sido generalizada, y puede decirse que la serie

$$a_0 + a_1x + a_2x^2 + \ldots$$

es asintótica a $f(x)$ en $x = 0$ si

$$\lim_{x\to 0} \frac{1}{x^n}\left[f(x) - \sum_0^{n-1} a_ix^i\right] = a_n.$$

A pesar de que en el caso de algunas series asintóticas se sabe qué error se comete al truncar por un determinado término, para las series asintóticas generales no se suele conocer tal información sobre el error numérico cometido. Sin embargo, las series asintóticas pueden ser utilizadas para dar resultados numéricos bastante aproximados para valores grandes de x usando sólo aquellos términos para los cuales el valor va disminuyendo según vayamos tomando más y más términos. El orden de magnitud del error es en cualquier caso el del primer término omitido.

Poincaré demostró que la suma, diferencia, producto y cociente de dos funciones vienen representadas asintóticamente por la suma, diferencia, producto y cociente de sus correspondientes series asintóticas, siempre que el término constante de la serie divisor no sea nulo. Por otra parte, si

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots,$$

entonces

$$\int_{x_0}^{x} f(z)\,dz \to C + a_0 x + a_1 \log x - \frac{a_2}{x} - \frac{1}{2}\frac{a_3}{x^2} -$$

El uso de la integración supone una ligera generalización de la definición original, a saber,

$$\phi(x) \sim f(x) + g(x)\left(a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots\right)$$

si

$$\frac{\phi(x) - f(x)}{g(x)} \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots$$

incluso cuando $f(x)$ y $g(x)$ no admitan ellas mismas representaciones por series asintóticas. En cuanto a la diferenciación, si se sabe que $f'(x)$ tiene un desarrollo en serie asintótica, entonces se puede obtener simplemente diferenciando la serie asintótica de $f(x)$.

Si una función dada tiene un desarrollo en serie asintótica entonces es único, pero el recíproco es falso porque, por ejemplo, $(1 + x)^{-1}$ y $(1 + e^{-x}) \cdot (1 + x)^{-1}$ tienen el mismo desarrollo asintótico.

Poincaré aplicó su teoría de las series asintóticas a las ecuaciones diferenciales, como hemos dicho, y pueden encontrarse muchos ejemplos de tal uso en su tratado de mecánica celeste *Les Méthodes nouvelles de la Mécanique Céleste.*[22] El tipo de ecuaciones estudiadas en su artículo de 1886 es

$$P_n(x)y^{(n)} + P_{n-1}(x)y^{(n-1)} + \ldots + P_0(x)y = 0 \qquad (17)$$

donde los $P_i(x)$ son polinomios en x. En realidad Poincaré sólo se ocupó del caso de segundo orden, pero el método puede aplicarse perfectamente al caso general (17).

Los únicos puntos singulares de la ecuación (17) son los ceros de $P_n(x)$ y $x = \infty$. Para un punto singular regular *(Stelle der Bestimmtheit)* existen expresiones convergentes para las integrales, dadas por Fuchs (cap. 29, sec. 5). Consideremos entonces un punto

[22] Vol. 2, cap. 8, 1893.

singular irregular; mediante una transformación lineal este punto puede ser desplazado a ∞, mientras que la ecuación sigue conservando su forma original. Si P_n es de grado p, la condición de que $x = \infty$ sea un punto singular regular implica que los grados de P_{n-1}, P_{n-2}, ..., P_0, sean como máximo p-1, p-2, ... p-n respectivamente. Para un punto singular irregular uno (o más) de estos grados tiene que ser mayor. Poincaré demostró que para una ecuación diferencial de la forma (17), donde los grados de los P_i no exceden del grado de P_n existen n series de la forma

$$e^{ax}x^{\alpha}\left(A_0 + \frac{A_1}{x} + \dots\right)$$

que satisfacen formalmente la ecuación diferencial. También demostró que a cada una de estas series corresponde una solución exacta en forma de una integral a la que la serie es asintótica.

Los resultados de Poincaré están incluidos en el siguiente teorema debido a Jacob Horn (1867-1946).[23] Horn estudia la ecuación

$$y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_n(x)y^n = 0, \tag{18}$$

donde los coeficientes son funciones racionales de x que se suponen desarrollables, para x positivo y suficientemente grande, en series convergentes o asintóticas del tipo

$$a_r(x) = x^{rk}\left[a_{r,0} + \frac{a_{r,1}}{x} + \frac{a_{r,2}}{x^2} + \dots\right], \qquad r = 1, 2, \dots, n,$$

siendo k algún entero positivo o 0. Si para la ecuación (18) anterior las raíces m_1, m_2, ..., m_n de su ecuación caracteristica, es decir, de la ecuación algebraica

$$m^n + a_{1,0}m^{n-1} + \dots + a_{n,0} = 0,$$

son distintas, entonces la ecuación (18) tiene n soluciones linealmente independientes y_1, y_2, ..., y_n desarrollables asintóticamente para valores positivos grandes de x, en la forma

[23] Acta Math., 24, 1901, 289-308.

$$y_r \sim e^{f_r(x)} x^{\varrho_r} \sum_{j=0}^{\infty} \frac{A_{r,j}}{x^j}, \qquad r = 1, 2, ..., n,$$

donde $f_r(x)$ es un polinomio de grado $k + 1$ en x, cuyo coeficiente de la mayor potencia en x es $m_r/(k + 1)$, mientras que ϱ_r y $A_{r,j}$ son constantes, con $A_{r,0} = 1$. Los resultados de Poincaré y Horn han sido extendidos a otros diversos tipos de ecuaciones diferenciales y generalizados para incluir los casos en que las raíces de la ecuación característica no sean necesariamente distintas.

La existencia, forma y otras características de la solución en forma de serie asintótica cuando en la ecuación (18) se permite que la variable independiente tome valores complejos, fueron estudiadas por primera vez por Horn[24]. George David Birkhoff (1884-1944), uno de los primeros grandes matemáticos americanos, obtuvo un resultado general para este caso.[25] En este artículo no considera Birkhoff la ecuación (18) sino el sistema más general

$$\frac{dy_i}{dx} = \sum_{j=1}^{n} a_{ij}(x) y_j, \qquad i = 1, 2, ..., n, \tag{19}$$

donde, para $|x| > R$ tenemos, para cada a_{ii}

$$a_{ij}(x) \sim a_{ij} x^q + a_{ij}^{(1)} x^{q-1} + ... + a_{ij}^{(q)} + a_{ij}^{(q+1)} \frac{1}{x} + ...$$

y para el que la ecuación característica en α

$$|a_{ij} - \delta_{ij} \alpha| = 0$$

tiene raíces distintas. Birkhoff obtuvo soluciones en series asintóticas para las y_i que se verifican en diversos sectores del plano complejo con vértices en $x = 0$.

Mientras que los desarrollos en serie tanto de Poincaré como de los restantes autores que acabamos de mencionar eran en potencias de la variable independiente, otras investigaciones sobre soluciones de ecuaciones diferenciales mediante series asintóticas volvieron a considerar el problema original de Liouville (sec. 2) con un paráme-

[24] Math. Ann., 50, 1898, 525-556.
[25] Amer. Math. Soc. Trans., 10, 1909, 436-470 = *Coll. Math. Papers*, 1, 201-235.

tro. Birkhoff mismo se ocupó del problema y obtuvo un resultado general.[26] Estudió la ecuación

$$\frac{d^n z}{dx^n} + \varrho a_{n-1}(x, \varrho)\frac{d^{n-1}z}{dx^{n-1}} + \ldots + \varrho^n a_0(x, \varrho)z = 0 \qquad (20)$$

para $|\varrho|$ grande y para x sobre el intervalo $[a,b]$. Las funciones $a_i(x,\varrho)$ se suponen analíticas en el parámetro complejo ϱ en $\varrho = \infty$ y con derivadas de todos los órdenes en la variable real x. Estas hipótesis sobre las $a_i(x,\varrho)$ implican que

$$a_i(x, \varrho) = \sum_{j=0}^{\infty} a_{ij}(x)\varrho^{-j}$$

y que las raíces de $w_1(x), w_2(x), \ldots, w_n(x)$ de la ecuación característica

$$w^n + a_{n-1,0}(x)w^{n-1} + \ldots + a_{00}(x) = 0$$

son distintas para cada x. Birkhoff demuestra que hay n soluciones independientes

$$z_1(x, \varrho), \ldots, z_n(x, \varrho)$$

de la ecuación (20) que son analíticas en ϱ en una región S del plano ϱ (determinada por el argumento de ϱ), tales que para todo m entero y para $|\varrho|$ grande

$$z_i(x, \varrho) = u_i(x, \varrho) + \exp\left[\varrho\int_a^x w_i(t)\,dt\right]E_0\varrho^{-m}, \qquad (21)$$

donde

$$u_i(x, \varrho) = \exp\left[\varrho\int_a^x w_i(t)\,dt \sum_{j=0}^{m-1} u_{ij}(x)\varrho^{-j}\right] \qquad (22)$$

y E_0 es una función de x,ϱ, y m acotada para todo x en $[a,b]$ y ϱ en S; las funciones $u_{ij}(x)$ pueden calcularse. A la vista de (22), el

[26] Amer. Math. Soc. Trans., 9, 1908, 219-231 y 380-382 = *Coll. Math. Papers*, 1, 1-36.

resultado (21) afirma que z_i viene dada por una serie en $1/\varrho$ hasta $1/\varrho^{m-1}$ más un término de resto, a saber, el segundo término de la derecha, que contiene $1/\varrho^m$. Además, dado que m es arbitrario, pueden tomarse en la expresión de $u_i(x,\varrho)$ tantos términos en $1/\varrho$ como se quiera. Y como E_0 está acotada, el término del resto es de orden más alto en $1/\varrho$ que $u_i(x,\varrho)$, y la serie completa que se puede obtener haciendo m infinito es asintótica a $z_i(x\varrho)/\exp\ (\varrho \int_a^x w\ (t)\ dt)$ en el sentido de Poincaré.

En el teorema de Birkhoff, la serie asintótica para ϱ compleja es válida solamente en un sector S del plano complejo ϱ. Aparece el fenómeno de Stokes, es decir, la prolongación analítica de $z_i(x,\varrho)$ a través de una línea de Stokes no viene dada por la prolongación analítica de la serie asintótica para $z_i(x,\varrho)$.

El uso de las series asintóticas o de la aproximación *WKBJ* a las soluciones de ecuaciones diferenciales planteó otro problema. Supongamos que tenemos la ecuación

$$y'' + \lambda^2 q(x)y = 0 \tag{23}$$

para x sobre el intervalo $[a,b]$. La aproximación *WKBJ* para valores grandes de λ da, en vista de (7), dos soluciones para $x > 0$ y dos para $x < 0$. Falla para los valores de x tales que $q = 0$. Uno de estos puntos recibe el nombre de punto de transición, punto crítico o punto de Stokes. Las soluciones exactas de la ecuación (23) son finitas en dichos puntos, sin embargo. El problema consiste en relacionar las soluciones *WKBJ* por cada lado del punto de transición de manera que representen la misma solución exacta sobre el intervalo $[a,b]$ sobre el que se trata de resolver la ecuación diferencial. Para formular el problema de una manera más explícita, considérese la ecuación anterior en la que $q(x)$ es real para x real y tal que $q(x) = 0$, $q'(x) \neq 0$ para $x = 0$. Supongamos también que $q(x)$ es negativa para x positiva (o al revés). Dada una combinación lineal de dos soluciones *WKBJ*, una para $x < 0$ y la otra para $x > 0$, se plantea la cuestión de cuál de las soluciones válidas para $x > 0$ habría que unir a una válida para $x < 0$. Las llamadas fórmulas de conexión dan la respuesta.

El primero que dio la idea de cómo cruzar a través de un cero de $q(x)$ fue Lord Rayleigh,[27] idea generalizada poco más tarde por

[27] Proc. roy. Soc., A86, 1912, 207-226 = *Sci. Papers*, 6, 71-90.

Richard Gans (1880-1954),[28] que conocía el trabajo de Rayleigh. Ambos eran físicos que estudiaban la propagación de la luz a través de medios no homogéneos.

El primer tratamiento sistemático de las fórmulas de conexión se debe a Harold Jeffreys,[29] independientemente del trabajo de Gans. Jeffreys estudió la ecuación

$$\frac{d^2y}{dx^2} + \lambda^2 X(x)y = 0, \qquad (24)$$

donde x es real, λ es real y suficientemente grande, y $X(x)$ tiene un cero simple, digamos en $x = 0$, obteniendo fórmulas para conectar las soluciones por medio de series asintóticas de (24) para $x > 0$ y $x < 0$ por medio de una ecuación de aproximación a (24), a saber

$$\frac{d^2y}{dx^2} + \lambda^2 xy = 0 \qquad (25)$$

en la que la $X(x)$ de (24) aparece sustituida por una función lineal en x. Las soluciones de (25) son

$$y_\pm(x) = x^{1/2} J_{\pm 1/3}(\xi) \qquad (26)$$

donde $\xi = (2/3)\, \lambda\, x^{3/2}$. Los desarrollos asintóticos de estas soluciones para valores grandes de x se pueden utilizar para unir las soluciones asintóticas de (24) por ambos lados de $x = 0$. Es necesario tener en cuenta diversos detalles sobre los dominios de x y λ en el proceso de unión, pero aquí no entraremos en ellos. Más tarde se publicaron numerosos artículos generalizando las fórmulas de conexión para los casos en que en (24) $X(x)$ tenga ceros múltiples o varios ceros distintos, así como para ecuaciones más complicadas de segundo orden y de orden superior, y para valores complejos de x y de λ.

La teoría de series asintóticas, aplicada al cálculo de integrales o a la resolución aproximada de ecuaciones diferenciales, ha experimentado un enorme crecimiento en años recientes. Vale la pena subrayar que los desarrollos matemáticos vienen a demostrar que los matemáticos del siglo XVIII y XIX, especialmente Euler, que se die-

---

[28] Annalen der Phys., (4), 47, 1915, 709-736.

[29] Proc. London Math. Soc., (2), 23, 1922-1924, 428-436.

ron cuenta de la gran utilidad de las series divergentes y sostuvieron que estas series se podrían usar como los equivalentes analíticos de las funciones que representaban, es decir, que operaciones sobre las series corresponderían a operaciones sobre las funciones, estaban bien encaminados. Aunque no consiguieran aislar rigurosamente la idea esencial, intuitivamente y a partir de los resultados obtenidos se dieron cuenta de que las series divergentes estaban estrechamente relacionadas con las funciones que representaban.

5. El problema de la sumabilidad de series divergentes

Los trabajos sobre series divergentes que hemos analizado hasta ahora se referían únicamente al problema de encontrar series asintóticas para representar funciones, o bien conocidas explícitamente, o cuya existencia estaba asegurada implícitamente como soluciones de ecuaciones diferenciales ordinarias. Otro problema que ocupó a los matemáticos desde 1880 aproximadamente en adelante, es esencialmente el recíproco del de encontrar series asintóticas. Dada una serie divergente en el sentido de Cauchy, ¿se le podrá asignar una «suma»? Si la serie es de términos variables, esta «suma» sería una función para la cual la serie divergente podría ser o no su desarrollo asintótico. No obstante, la función podría ser considerada como la «suma» de la serie, «suma» que podría servir para algún fin útil, incluso aunque la serie no convergiese a ella o no pudiera ser utilizable para calcular valores aproximados de la función.

En cierto sentido, el problema de sumar series divergentes se planteó ya antes de que Cauchy introdujese sus definiciones de convergencia y divergencia. Los matemáticos se encontraron con series divergentes y trataron de encontrar sumas para ellas lo mismo que para las series convergentes, simplemente porque la distinción entre los dos tipos no estaba establecida con precisión, y la única pregunta era ¿cuál será la suma correcta? Así, por ejemplo, el principio de Euler (cap. 20, sec. 7) de que un desarrollo en serie de potencias de una función tiene como suma el valor de la función de la que se deriva la serie, daba una suma para la serie incluso para valores de x para la que ésta diverge en el sentido de Cauchy. Análogamente, en su transformación de series (cap. 20, sec. 4) convertía las series divergentes en convergentes sin dudar de que prácticamente todas las series tendrían su suma. Sin embargo, a partir de que Cauchy

hiciese la distinción entre convergencia y divergencia, el problema de sumar series divergentes se planteaba en un nivel distinto. La asignación de sumas a todas las series, hecha de manera relativamente ingenua durante el siglo XVIII, ya no era aceptable. Las nuevas definiciones deberían establecer lo que hoy se llama *sumabilidad* para distinguirlo del concepto de convergencia en el sentido de Cauchy.

Mirando retrospectivamente, puede verse que la noción de sumabilidad era realmente la que estaban anticipando los matemáticos del siglo XVIII y comienzos del XIX. A esto es a lo que conducen los métodos de sumación de Euler que acabamos de mencionar; de hecho, en una carta a Goldbach de 7 de agosto de 1745, en la que afirma que la suma de una serie de potencias es el valor de la función de la que se deriva la serie, Euler asegura también que toda serie debe tener una suma, pero debido a que la palabra suma implica el proceso usual de añadir, y este proceso no conduce a la suma en el caso de una serie divergente tal como la $1 - 1! + 2! - 3! + \ldots$, tendríamos que utilizar la palabra «valor» para la «suma» de una serie divergente.

Poisson también introdujo lo que hoy reconoceríamos como un concepto de sumabilidad. En la definición de Euler de suma como el valor de la función de la que se deriva la serie, está implícita la idea de que

$$\sum_{n=0}^{\infty} a_n = \lim_{x\to 1-} \sum_{n=0}^{\infty} a_n x^n, \tag{27}$$

donde la notación $x \to 1-$ significa que x tiende a 1 por valores inferiores. Según (27) la suma de $1 - 1 + 1 - 1 + \ldots$ ha de ser

$$\lim_{x\to 1-} (1 - x + x^2 - x^3 + \ldots) = \lim_{x\to 1-} (1 + x)^{-1} = \frac{1}{2}$$

Poisson [30] estudió la serie

$$\text{sen}\ \theta + \text{sen}\ 2\theta + \text{sen}\ 3\theta + \ldots$$

[30] Jour. de l'Ecole Poly., 11, 1820, 417-489.

que diverge excepto cuando θ es un múltiplo de π. Su idea, expresada para la serie de Fourier completa

$$\frac{a_0}{2} + \sum_n (a_n \cos n\theta + b_n \operatorname{sen} n\theta) \tag{28}$$

es la de que habría que considerar la serie de potencias asociada

$$\frac{a_0}{2} + \sum_n (a_n \cos n\theta + b_n \operatorname{sen} n\theta)r^n \tag{29}$$

y definir la suma de (28) como el límite de la serie (29) cuando r tiende a 1 por la izquierda. Por supuesto, Poisson no era consciente de que lo que estaba sugiriendo era una definición de suma para una serie divergente porque, como ya hemos dicho, la distinción entre convergencia y divergencia no era clara en esta época.

La definición utilizada por Poisson recibe hoy el nombre de sumabilidad en el sentido de Abel, porque también aparece sugerida en un teorema de Abel [31] que afirma que si la serie de potencias

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

tiene radio de convergencia r y converge para $x = r$ entonces

$$\lim_{x\to r-} f(x) = \sum_{n=0}^{\infty} a_n r^n. \tag{30}$$

Entonces la función $f(x)$ definida por la serie en el intervalo $-r < x \leqslant r$ es continua por la izquierda en $x = r$. Sin embargo, si Σa_n no converge y el límite (30) existe para $r = 1$, entonces tenemos una definición de suma para la serie divergente. Esta definición formal de sumabilidad para series divergentes en el sentido de Cauchy no se introdujo hasta finales del siglo XIX en un contexto que veremos más adelante.

Uno de los motivos para reconsiderar la sumación de series divergentes, aparte de su reconocida utilidad en astronomía, fue lo que se ha llamado el problema del valor en la frontera *(Grenzwert)* en

[31] Jour. für Math., 1, 1826, 311-339 = *Œuvres*, 1, 219-250.

la teoría de funciones analíticas. Una serie de potencias $\Sigma a_n x^n$ puede representar una función analítica dentro de un círculo de radio r pero no para valores de x en la circunferencia. El problema era saber si uno podría encontrar un concepto de suma tal que la serie de potencias pudiera tener una suma para $|x| = r$ y tal que dicha suma pudiera incluso ser el valor de $f(x)$ cuando $|x|$ tiende a r. Este intento de extender el dominio de representación de una función analítica mediante una serie de potencias fue lo que motivó a Frobenius, Hölder y Ernesto Cesàro. Frobenius [32] demostró que si la serie de potencias $\Sigma a_n x^n$ tiene el intervalo de convergencia $-1 < x < 1$, y si

$$s_n = a_1 + a_2 + a_3 + \ldots + a_n. \tag{31}$$

entonces

$$\lim_{x\to 1-} \sum_{n=0}^{\infty} a_n x^n = \lim_{n\to\infty} \frac{s_0 + s_1 + \ldots + s_n}{n+1}$$

cuando existe el límite de la derecha. Así, la serie de potencias, normalmente divergente para $x = 1$, puede tener una suma; además, si $f(x)$ es la función representada por la serie, la definición de Frobenius del valor de la serie para $x = 1$ coincide con $\lim_{x\to 1-} f(x)$.

Abstrayéndolo de su relación con las series de potencias, el artículo de Frobenius venía a sugerir una definición de sumabilidad para series divergentes. Si Σa_n es divergente y s_n tiene el significado de (31), entonces se puede tomar como suma

$$\sum_{n=0}^{\infty} a_n = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{s_0 + s_1 + \ldots + s_n}{n+1}$$

si este límite existe. Por ejemplo, para la serie $1 - 1 + 1 - 1 + \ldots$, las S_n toman los valores 1, 1/2, 2/3, 2/4, 3/5, 1/2, 4/7, 1/2, ... de manera que $\lim_{n\to\infty} S_n = 1/2$. Si $\Sigma\, a_n$ converge, entonces la «suma» de Frobenius da la suma usual. Esta idea de promediar las sumas parciales de una serie se puede encontrar ya en la vieja literatura; fue utilizada para tipos especiales de series por Daniel Bernoulli [33] y Joseph L. Raabe (1801-1859).[34]

---

[32] Jour. für Math., 89, 1880, 262-264 = *Ges. Abh.*, 2, 8-10.
[33] Comm. Acad. Sci. Petrop. 16, 1771, 71-90.
[34] Jour. für Math., 15, 1836, 355-364.

Poco después de que Frobenius publicara su artículo, Hölder dio una generalización. Dada la serie Σa_n sean

$$s_n^{(0)} = s_n$$

$$s_n^{(1)} = \frac{1}{n+1}(s_0^{(0)} + s_1^{(0)} + \dots + s_n^{(0)})$$

$$s_n^{(2)} = \frac{1}{n+1}(s_0^{(1)} + s_1^{(1)} + \dots + s_n^{(1)})$$

....

$$s_n^{(r)} = \frac{1}{n+1}(s_0^{(r-1)} + s_1^{(r-1)} + \dots + s_n^{(r-1)}).$$

Entonces la suma s viene dada por

$$s = \lim_{n\to\infty} s_n^{(r)} \tag{32}$$

si este límite existe para algún r. La definición de Hölder se conoce hoy como (H,r)-sumabilidad.

He aquí un ejemplo de Hölder: considérese la serie

$$-\frac{1}{(1+x)^2} = -1 + 2x - 3x^2 + 4x^3 - \dots$$

que diverge para $x = 1$. Sin embargo, para $x = 1$ se tiene

$$s_0 = -1, s_1 = 1, s_2 = -2, s_3 = 2, s_4 = -3, \dots$$

Entonces

$$s_0{}^{(1)} = -1, s_1^{(1)} = 0, s_2^{(1)} = -\frac{2}{3}, s_3^{(1)} = 0, s_4^{(1)} = -\frac{3}{5}, \dots;$$

$$s_0^{(2)} = -1, s_1^{(2)} = -\frac{1}{2}, s_2^{(2)} = -\frac{5}{9}, s_3^{(2)} = -\frac{5}{12}, s_4^{(2)} = -\frac{34}{75}, \dots .$$

Es casi inmediato que $\lim_{n\to\infty} s_n{}^{(2)} = -1/4$, y esta es la $(H,2)$-suma de Hölder. Se trata exactamente del mismo valor que asignó

---

[35] Math. Ann., 20, 1882, 535-549.

Euler a la serie basándose en su principio de que la suma es el valor de la función de la que se deriva la serie.

Otra de las definiciones de sumabilidad hoy habituales se debe a Cesàro, profesor de la universidad de Nápoles.[36] Sea la serie $\Sigma_{i=0}^{\infty} a_i$ y sea $s_n = \Sigma_{i=0}^{n} a_i$. Entonces la suma de Cesàro es

$$s = \lim_{n\to\infty} \frac{S_n^{(r)}}{D_n^{(r)}}, \qquad r \text{ entero y } \geq 0, \tag{33}$$

donde

$$S_n^{(r)} = s_n + rs_{n-1} + \frac{r(r+1)}{2!} s_{n-2} + \ldots + \frac{r(r+1)\ldots(r+n-1)}{n!} s_0$$

y

$$D_n^{(r)} = \frac{(r+1)(r+2)\ldots(r+n)}{n!}$$

El caso $r = 1$ incluye la definición de Frobenius. La definición de Cesàro se conoce hoy como (C,r)-sumabilidad. Los métodos de Hölder y Cesàro dan los mismos resultados: que la sumabilidad de Hölder implica la de Cesàro fue demostrado por Konrad Knopp (1882-1957) en una disertación no publicada en 1907; el recíproco fue demostrado por Walter Schnee (n. 1885).[37]

Una característica interesante de algunas de las definiciones de sumabilidad, cuando se aplican a series de potencias de radio de convergencia 1, es la de que no sólo dan una suma que coincide con $\lim_{x \to 1^-} f(x)$, donde $f(x)$ es la función de la que se deriva la serie, sino que tienen la propiedad adicional de que siguen teniendo significado en regiones donde $|x| > 1$ y en estas regiones dan la prolongación analítica de la serie de potencias original.

Otros desarrollos en la sumación de series divergentes estuvieron motivados por un problema completamente distinto, las investigaciones de Stieltjes sobre fracciones continuas. Euler [38] había utiliza-

---

[36] Bull. Sci. Math., (2), 1890, 114-120.

[37] Math. Ann., 67, 1909, 110-125.

[38] Novi Comm. Acad. Sci. Petrop.,5, 1754-1755, 205-237, pub. 1760 = *Opera*. (1), 14, 585-617, y Nova Acta Acad. Sci. Petrop., 2, 1784, 36-45, pub. 1788 = *Opera*, (1), 16, 34-43.

do ya el hecho de que las fracciones continuas pueden convertirse en series convergentes o divergentes y recíprocamente. Euler buscaba (cap. 20, sec. 4 y 6) una suma para la serie divergente

$$1 - 2! + 3! - 4! + 5! - \dots . \tag{34}$$

En primer lugar demostró, en su artículo sobre series divergentes [39] y en cartas a Nicholas Bernoulli (1687-1759),[40] que

$$x - (1!)x^2 + (2!)x^3 - (3!)x^4 + \dots \tag{35}$$

satisface formalmente la ecuación diferencial

$$x^2 \frac{dy}{dx} + y = x$$

para la que obtuvo la solución integral

$$y = \int_0^\infty \frac{xe^{-t}}{1 + xt}\, dt. \tag{36}$$

Entonces, utilizando reglas obtenidas transformando series convergentes en fracciones continuas, Euler transformó (35) en

$$\frac{x}{1+}\,\frac{x}{1+}\,\frac{x}{1+}\,\frac{2x}{1+}\,\frac{2x}{1+}\,\frac{3x}{1+}\,\frac{3x}{1+}\dots . \tag{37}$$

Este trabajo tiene dos características importantes. Por un lado Euler obtiene una integral que puede ser considerada como la «suma» de la serie divergente (35); de hecho, esta es asintótica a la integral. Por otra parte mostró cómo convertir series divergentes en fracciones continuas. Utilizó la fracción continua para $x = 1$ para calcular el valor de la serie (34).

Otros trabajos de este tipo aparecieron a finales del siglo XVIII y buena parte del XIX, de los que el más notable se debe a Laguerre,[41]

---

[39] Novi Comm. Acad. Sci. Petrop., 5, 1754-1755, 205-237, pub. 1760 = *Opera*, (1), 14, 585-617.

[40] Euler, *Opera Posthuma*, 1, 545-549.

[41] Bull. Soc. Math. de France, 7, 1879, 72-81 = *Œuvres*,, 1, 428-437.

que el cociente C_n/C_{n-1} aumenta; si tiene un límite finito λ, la serie converge para $|z| > \lambda$, pero si el cociente crece indefinidamente la serie diverge para todo z.

La relación entre la serie (42) y la fracción continua (40) es más complicada. Aunque la fracción continua converge si lo hace la serie, el recíproco no es cierto. Cuando la serie (42) diverge hay que distinguir dos casos, según que Σa_n sea divergente o convergente. En el primer caso, como hemos dicho, la fracción continua da uno y sólo un equivalente funcional, que puede tomarse como la suma de la serie divergente (42). Cuando Σa_n es convergente se obtienen dos funciones distintas de la fracción continua, una de los convergentes pares y la otra de los convergentes impares. Pero a la serie (42) (ahora divergente) corresponden infinitas funciones, cada una de las cuales tiene dicha serie como desarrollo asintótico.

Los resultados de Stieltjes también tuvieron la importante consecuencia siguiente: revelaban una clasificación de las series divergentes en dos clases al menos; aquellas series para las que propiamente había un único equivalente funcional cuyo desarrollo era dicha serie, y aquellas para las que había al menos dos equivalentes funcionales cuyo desarrollo era la serie. La fracción continua es solamente el intermediario entre la serie y la integral; es decir, dada la serie, se obtiene la integral a través de la fracción continua. Así, una serie divergente corresponde a una o más funciones, las cuales pueden considerarse como la suma de la serie en un sentido nuevo de la misma.

Stieltjes también planteó y resolvió un tipo de problema inverso. Para simplificar un poco la exposición, supongamos que $\phi(u)$ es diferenciable, de manera que la integral (41) se puede escribir como

$$\int_0^\infty \frac{f(u)}{z+u}\,du.$$

A la serie divergente (42), y en el caso en que $\Sigma\, a_n$ sea divergente, corresponde una integral de este tipo. El problema recíproco es, dada la serie, hallar $f(u)$. Un desarrollo formal de la integral muestra que

$$C_n = \int_0^\infty f(u)u^n du \qquad n = 0, 1, 2, \ldots \,. \tag{43}$$

en 1894 y 1895.[43] Esta obra, que constituye el origen de la teoría analítica de fracciones continuas, estudia cuestiones de convergencia y las relaciones con las integrales definidas y series divergentes. En estos artículos fue donde introdujo Stieltjes la integral que lleva su nombre.

Stieltjes parte de la fracción continua

$$\frac{1}{a_1 z+}\ \frac{1}{a_2+}\ \frac{1}{a_3 z+}\ \frac{1}{a_4+}\ \frac{1}{a_5 z+}\cdots\frac{1}{a_{2n}+}\ \frac{1}{a_{2n+1} z+}\cdots, \tag{40}$$

donde los a_n son números reales positivos y z es una variable compleja. Demuestra que cuando la serie $\sum_{n=1}^{\infty} a_n$ diverge, la fracción continua (40) converge a una función $F(z)$ que es analítica en todo el plano complejo excepto a lo largo del eje real negativo junto con el origen, y

$$F(z) = \int_0^\infty \frac{d\phi(u)}{u+z}. \tag{41}$$

Cuando Σa_n converge, las sumas parciales pares e impares de (40) convergen a límites distintos $F_1(z)$ y $F_2(z)$, donde

$$F_1(z) = \int_0^\infty \frac{dg_1(u)}{z+u}, \qquad F_2(z) = \int_0^\infty \frac{dg_2(u)}{z+u}.$$

Por otra parte, se sabía que la fracción continua (40) podía desarrollarse formalmente en serie

$$\frac{C_0}{z} - \frac{C_1}{z^2} + \frac{C_2}{z^3} - \frac{C_3}{z^4} + \cdots \tag{42}$$

con los C_i positivos. La correspondencia es también recíproca (con algunas restricciones); a toda serie (42) le corresponde una fracción continua (40) con los a_n positivos. Stieltjes mostró cómo calcular los C_i a partir de los a_n, y en el caso de que Σa_n sea divergente, demostró

[43] Ann. Fac. Sci. de Toulouse, 8, 1894, J. 1-122, y 9, 1895,A. 1-47 = *Œuvres complètes*, 2, 402-559.

gación analítica de la serie original. Borel llama a la serie *sumable* (en el sentido mencionado) en un punto x donde la integral tenga sentido.

Si la serie original (44) es divergente ($R = 0$), la serie asociada puede ser convergente o divergente. Si es convergente sobre una región del plano $u = zx$ únicamente, entendemos por $F(u)$ el valor no simplemente de la serie asociada, sino el de su prolongación analítica. Entonces la integral $\int_0^\infty e^{-z} F(zx)\, dz$ puede tener un significado y se obtiene, como se ve, de la serie divergente original. Borel se ocupó también de la determinación de la región de los valores de x en los que la serie original es sumable, tanto cuando la serie original es convergente ($R > 0$) como si es divergente ($R = 0$).

Borel introdujo también el concepto de sumabilidad absoluta. La serie original se llama absolutamente sumable en un valor de x cuando

$$\int_0^\infty e^{-z}F(zx)dz$$

es absolutamente convergente y las sucesivas integrales

$$\int_0^\infty e^{-z}\left|\frac{d^\lambda F(zx)}{dz^\lambda}\right| dz, \qquad \lambda = 1, 2, \ldots$$

tengan sentido. Borel demuestra a continuación que si una serie divergente es absolutamente sumable, puede ser manipulada precisamente como una serie convergente. En otras palabras, la serie representa una función y puede ser utilizada en lugar de la función.

Así, la suma, diferencia y producto de dos series absolutamente sumables es absolutamente sumable y representa la suma, diferencia y producto, respectivamente, de las dos funciones representadas por las series individuales. Un hecho análogo se verifica para la derivada de una serie absolutamente sumable. Además, la suma en el sentido anterior coincide con la suma usual en el caso de las series convergentes, y la resta de los k primeros términos reduce la «suma» de la serie total en la suma de esos k términos. Borel subraya que cualquier definición satisfactoria de sumabilidad debería tener estas propiedades, aunque no todas las tienen. No exige, en cambio, que dos definiciones cualesquiera den necesariamente la misma suma.

Por tanto, conociendo los C_n hay que determinar $f(u)$ de manera que satisfaga las infinitas ecuaciones (43). A este problema es al que llamó Stieltjes «problema de los momentos». No admite solución única, y Stieltjes mismo dio una función

$$f(u) = e^{-\sqrt[4]{u}} \operatorname{sen} \sqrt[4]{u}$$

que hace $C_n = 0$ para todo n. Si se impone la condición suplementaria de que $f(u)$ sea positiva entre los límites de integración, entonces sólo es posible una única solución $f(u)$.

El desarrollo sistemático de la teoría de series sumables comienza con la obra de Borel a partir de 1895. Al principio Borel dio definiciones que generalizan las de Cesàro, para dar después, partiendo de la obra de Stieltjes, una definición integral.[44] Si se aplica el proceso utilizado por Laguerre a una serie cualquiera de la forma

$$a_0 + a_1x + a_2x^2 + \ldots \tag{44}$$

que tenga un radio de convergencia finito (incluyendo 0), nos vemos conducidos a la integral

$$\int_0^\infty e^{-z}F(zx)dz, \tag{45}$$

donde

$$F(u) = 1 + a_1u + \frac{a_2}{2!}u^2 + \ldots + \frac{a_n}{n!}u^n + \ldots .$$

Esta integral es la expresión sobre la que construyó Borel su teoría de series divergentes; Borel la consideró como la suma de la serie (44). La serie $F(u)$ se llama la serie asociada de la serie original.

Si la serie original (44) tiene un radio de convergencia R mayor que cero, entonces la serie asociada representa una función entera, la integral $\int_0^\infty e^{-z} F(zx)\,dz$ tiene sentido si x es interior al círculo de convergencia y los valores da la integral y la serie son idénticos. Pero la integral también puede tener sentido para valores de x exteriores al círculo de convergencia y en este caso la integral da una prolon-

[44] Ann. de l'Ecole Norm. Sup., (3), 16, 1899, 9-136.

gación analítica de la serie original. Borel llama a la serie *sumable* (en el sentido mencionado) en un punto x donde la integral tenga sentido.

Si la serie original (44) es divergente ($R = 0$), la serie asociada puede ser convergente o divergente. Si es convergente sobre una región del plano $u = zx$ únicamente, entendemos por $F(u)$ el valor no simplemente de la serie asociada, sino el de su prolongación analítica. Entonces la integral $\int_0^\infty e^{-z} F(zx)\, dz$ puede tener un significado y se obtiene, como se ve, de la serie divergente original. Borel se ocupó también de la determinación de la región de los valores de x en los que la serie original es sumable, tanto cuando la serie original es convergente ($R > 0$) como si es divergente ($R = 0$).

Borel introdujo también el concepto de sumabilidad absoluta. La serie original se llama absolutamente sumable en un valor de x cuando

$$\int_0^\infty e^{-z}F(zx)dz$$

es absolutamente convergente y las sucesivas integrales

$$\int_0^\infty e^{-z} \left|\frac{d^\lambda F(zx)}{dz^\lambda}\right| dz, \qquad \lambda = 1, 2, \ldots$$

tengan sentido. Borel demuestra a continuación que si una serie divergente es absolutamente sumable, puede ser manipulada precisamente como una serie convergente. En otras palabras, la serie representa una función y puede ser utilizada en lugar de la función.

Así, la suma, diferencia y producto de dos series absolutamente sumables es absolutamente sumable y representa la suma, diferencia y producto, respectivamente, de las dos funciones representadas por las series individuales. Un hecho análogo se verifica para la derivada de una serie absolutamente sumable. Además, la suma en el sentido anterior coincide con la suma usual en el caso de las series convergentes, y la resta de los k primeros términos reduce la «suma» de la serie total en la suma de esos k términos. Borel subraya que cualquier definición satisfactoria de sumabilidad debería tener estas propiedades, aunque no todas las tienen. No exige, en cambio, que dos definiciones cualesquiera den necesariamente la misma suma.

Estas propiedades posibilitaron la aplicación inmediata de la teoría de Borel a las ecuaciones diferenciales. De hecho, si

$$P(x, y, y', ..., y^{(n)}) = 0$$

es una ecuación diferencial que es holomorfa en x en el origen y algebraica en y y todas sus derivadas, entonces cualquier serie absolutamente sumable

$$a_0 + a_1x + a_2x^2 + ...$$

que satisfaga formalmente la ecuación diferencial, define una función analítica que es una solución de la ecuación. Por ejemplo, la serie de Laguerre

$$1 + x + 2!x^2 + 3!x^3 + ...$$

satisface formalmente la ecuación

$$x^2 \frac{d^2y}{dx^2} + (x - 1)y = -1,$$

y por tanto la función (39)

$$f(x) = \int_0^\infty e^{-z} \frac{1}{1 - zx} dz$$

será solución de la ecuación.

Una vez que el concepto de sumabilidad fue aceptado, docenas de matemáticos introdujeron diversas definiciones nuevas que cumplían algunas o todas las exigencias indicadas por Borel y otros. Muchas de las definiciones de sumabilidad han sido generalizadas a series dobles. Naturalmente también se han planteado y resuelto muchos problemas que involucran la noción de sumabilidad. Por ejemplo, supongamos una serie sumable por un cierto método: ¿qué condiciones adicionales habría que imponer a dicha serie para que, supuesta su sumabilidad, también fuera convergente en el sentido de Cauchy? Tales teoremas reciben el nombre de teoremas tauberianos, en honor a Alfred Tauber (n. 1866). Así, por ejemplo, Tauber de-

mostró [45] que si Σa_n es sumable en el sentido de Abel a s y na_n tiende a cero cuando n tiende a infinito, entonces Σa_n converge a s.

El concepto de sumabilidad nos permite, pues, dar un valor o suma a una gran variedad de series divergentes. Naturalmente, se plantea de manera necesaria la cuestión de qué es lo que se ha conseguido así. Si de una situación física concreta surge una serie dada, la adecuación de cualquier definición de suma dependería completamente de si tal suma es o no físicamente significativa, de la misma manera que la utilidad física de cualquier geometría depende de si describe o no adecuadamente el espacio físico. La definición de suma de Cauchy es la que usualmente encaja bien, porque básicamente viene a decir que la suma es lo que se obtiene añadiendo cada vez más términos en el sentido ordinario. Pero no hay ninguna razón lógica para preferir este concepto de suma a los otros que han sido introducidos. Ciertamente, la representación de funciones mediante series se ha ampliado enormemente utilizando los nuevos conceptos. Por ejemplo, Leopold Fejér (1880-1959), alumno de H. A. Schwarz, mostró el valor de la teoría de sumabilidad dentro de la teoría de las series de Fourier. En 1904 [46] demostró Fejér que si la función $f(x)$ está acotada en el intervalo $[-\pi,\pi]$ y es integrable en el sentido de Riemann, o si no está acotada pero la integral $\int \pi\ f(x)\ dx$ es absolutamente convergente, entonces en todo punto del intervalo en el que existan $f(x + 0)$ y $f(x - 0)$, la suma de Frobenius de la serie de Fourier

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

es $[f(x + 0) + f(x - 0)]/2$. Las condiciones exigidas a $f(x)$ en este teorema son más débiles que en los teoremas anteriores sobre la convergencia de la serie de Fourier a $f(x)$ (véase cap. 40, sec. 6).

Este resultado fundamental de Fejér fue el comienzo de una larga serie de investigaciones fructíferas sobre la sumabilidad de series. Ya hemos tenido numerosas ocasiones de ver la necesidad de representar las funciones por medio de series. Así, al tratar la condición inicial en la resolución de problemas de valor inicial o de contorno

---

[45] Monatshefte fur Mathematik und Physik, 8, 1897, 273-277.

[46] Math. Ann., 58, 1904, 51-69.

para ecuaciones en derivadas parciales, suele ser necesario representar la $f(x)$ inicial dada en términos de autofunciones obtenidas aplicando las condiciones de contorno a las ecuaciones diferenciales ordinarias que resultan del método de separación de variables. Estas autofunciones pueden ser funciones de Bessel, funciones de Legendre o cualesquiera otras de un cierto número de funciones especiales. Mientras que puede no tenerse la convergencia en el sentido de Cauchy de tal serie de autofunciones a la $f(x)$ dada, dicha serie puede ciertamente ser sumable a $f(x)$ en uno u otro de los sentidos de sumabilidad que hemos visto y así se satisfaga la condición inicial. Estas aplicaciones de la sumabilidad representan un gran éxito del concepto mismo.

La construcción y posterior aceptación de la teoría de series divergentes constituye otro ejemplo sorprendente de cómo crece la matemática. Viene a mostrar, ante todo, que cuando un concepto o técnica demuestra ser útil, incluso aunque su lógica sea confusa o inexistente, una investigación persistente descubrirá alguna justificación lógica a propósito. También muestra lo lejos que han llegado los matemáticos en reconocer que la matemática es obra del hombre. Las nuevas definiciones de sumabilidad no coinciden con la noción natural de ir añadiendo más y más términos continuamente, noción que Cauchy simplemente rigorizó, sino que son artificiales. Pero no obstante sirven para fines matemáticos, incluyendo la solución matemática de problemas físicos; y estos son ahora motivos suficientes para admitirlas dentro del dominio de la matemática con todos los derechos.

Bibliografía

Borel, Emile: *Notice sur les travaux scientifiques de M. Emile Borel*, 2.ª ed., Gauthier-Villars, 1921.

— *Leçons sur les séries divergentes*, Gauthier-Villars, 1901.

Burkhardt, H.: «Trigonometrische Reihe und Integrale». *Encyk. der Math. Wiss.*, B. G. Teubner, 1904-1916, II A12, 819-1354.

— «Uber den Gebrauch divergenter Reihen in der Zeit von 1750-1860». *Math. Ann.*, 70, 1911, 169-206.

Carmichael, Robert D.: «General Aspects of the Theory of Summable Series». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 97-131.

Collingwood, E. F.: «Emile Borel». *Jour. Lon. Math. Soc.*, 34, 1959, 488-512.

Ford, W. B.: «A Conspectus of the Modern Theories of Divergent Series». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 1-15.

Hardy, G. H.: *Divergent Series*, Oxford University Press, 1949. Véanse las notas históricas.

Hurwitz, W. A.: «A Report on Topics in the Theory of Divergent Series». *Amer. Math. Soc. Bull.*, 28, 1922, 17-36.

Knopp, K.: «Neuere Untersuchungen in der Theorie der divergenten Reihen». *Jahres. der Deut. Math.-Verein.*, 32, 1923, 43-67.

Langer, Rudolf E.: «The Asymptotic Solution of Ordinary Linear Differential Equations of the Second Order». *Amer. Math. Soc. Bull.*, 40, 1934, 545-582.

McHugh, J. A. M.: «An Historical Survey of Ordinary Linear Differential Equations with a Large Parameter and Turning Points». *Archive for History of Exact Sciences*, 7, 1971, 277-324.

Moore, C. N.: «Applications of the Theory of Summability to Developments in Orthogonal Functions». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 258-276.

Plancherel, Michel: «Le Développement de la théorie des séries trigonométriques dans le dernier quart de siècle». *L'Enseignement Mathématique*, 24, 1924/1925, 19-58.

Pringsheim, A.: «Irrationalzahlen und Konvergenz unendlicher Prozesse». *Encyk. der Math. Wiss.*, B. G. Teubner, 1898-1904, IA3, 47-146.

Reiff, R.: *Geschichte der unendlichen Reihen*, H. Lauppsche Buchhandlung, 1889, Martin Sandig (reprint), 1969.

Smail, L. L.: *History and Synopsis of the Theory of Summable Infinite Processes*, University of Oregon Press, 1925.

Van Vleck, E. B.: «Selected Topics in the Theory of Divergent Series and Continued Fractions». *The Boston Colloquium of the Amer. Math. Soc.*, 1903, Macmillan, 1905, 75-187.

Capítulo 48

EL ANALISIS TENSORIAL Y LA GEOMETRIA DIFERENCIAL

> Por tanto, o bien la realidad que sirve de base al espacio debe consistir en una variedad discreta, o bien tendremos que buscar el fundamento de su métrica en relaciones exteriores a ella, en las fuerzas de unión que actúan sobre ella. Esto nos conduce a los dominios de otra ciencia, los de la física, a los que el objetivo de nuestro trabajo no nos permite ir ahora.
>
> BERNHARD RIEMANN

1. Los orígenes del análisis tensorial

A menudo se habla del análisis tensorial como si se tratase de una rama totalmente nueva de la matemática, creada *ab initio*, bien para satisfacer algún objetivo concreto, o para deleite de los matemáticos. En realidad se trata solamente de una nueva variación sobre un viejo tema, a saber, el del estudio de los invariantes diferenciales asociados en principio a una geometría riemanniana. Recordemos (vid. cap. 37, sec. 5) que esos invariantes eran expresiones que conservaban su forma y su valor bajo cualquier cambio en el sistema de coordenadas, porque representaban propiedades puramente geométricas o físicas.

El estudio de los invariantes diferenciales había sido iniciado por Riemann, Beltrami, Christoffel y Lipschitz. El nuevo enfoque se debe a Gregorio Ricci-Curbastro (1853-1925), que fue profesor de matemáticas en la Universidad de Padua. Influyó mucho sobre él Luigi Bianchi, cuya obra había continuado la de Christoffel. Ricci trató de facilitar la búsqueda de propiedades geométricas y la expresión de leyes físicas en forma invariante bajo cambios del sistema de coordenadas. Realizó sus trabajos más importantes sobre el tema

durante los años 1887-1896, aunque tanto él mismo como el resto de la escuela italiana continuaron trabajando en ello durante veinte años o más a partir de 1896. En esos primeros nueve años Ricci desarrolló sus planteamientos y elaboró un sistema de notación completo para la teoría, que llamó «cálculo difrencial absoluto». Presentó la primera exposición sistemática de su método en un artículo publicado en 1892,[1] aplicándolo a algunos problemas de geometría diferencial y de física.

Nueve años más tarde, y en colaboración con su famoso discípulo Tullio Levi-Civita (1873-1941), publicó un artículo general, «Métodos de cálculo diferencial absoluto», ofreciendo una formulación más definitiva de ese cálculo. El tema terminaría por conocerse como «análisis tensorial» desde que Einstein le diera ese nombre en 1916. Dados los muchos cambios de notación realizados por Ricci, y más tarde por Levi-Civita y Ricci, utilizaremos la notación aceptada actualmente de una manera bastante general.

2. El concepto de tensor

Para llegar a la idea de tensor tal como la introdujo Ricci, consideremos una función $A(x^1, x^2, ..., x^n)$, cuyas derivadas parciales $\partial A/\partial x^i$ representaremos por A_i. Entonces la expresión

$$\sum A_j \, dx^j \tag{1}$$

es un invariante diferencial bajo transformaciones del tipo

$$x^i = f_i(y^1, y^2, ..., y^n). \tag{2}$$

Supondremos que las funciones f_i tienen todas las derivadas necesarias y que la transformación es reversible, de manera que

$$y^i = g_i(x^1, x^2, ..., x^n). \tag{3}$$

[1] Bull. des Sci. Math., (2), 16, 1892, 167-189 = *Opere*, 1, 288-310.

[2] Math. Ann., 54, 1901, 125-201 = Ricci, *Opere*, 2, 185-271.

Bajo la transformación (2) la expresión (1) se convierte en

$$\sum \bar{A}_j(y^1, y^2, ..., y^n)dy^j. \qquad (4)$$

Sin embargo, $\bar{A}_j$ no será en general igual a A_j, sino que

$$\bar{A}_j = \frac{\partial \bar{A}}{\partial y^j} = A_1 \frac{\partial x^1}{\partial y^j} + A_2 \frac{\partial x^2}{\partial y^j} + A_2 \frac{\partial x^2}{\partial y^j} + ... + A_n \frac{\partial x^n}{\partial y^j}, \qquad (5)$$

donde se entiende que las x^i han sido sustituidas en las A_i por sus valores en términos de las y^i. Así pues, las $\bar{A}_j$ se relacionan con las A_j por medio de la ley de transformación (5), en la que intervienen las derivadas primeras de las ecuaciones de la transformación.

La idea básica de Ricci era que, en lugar de referirse a la forma diferencial invariante (1), sería suficiente, y más sencillo, considerar simplemente el conjunto de funciones

$$A_1, A_2, ..., A_n$$

y llamarlas componentes de un tensor siempre que, bajo un cambio de coordenadas, el nuevo conjunto de componentes

$$\bar{A}_1, \bar{A}_2, ..., \bar{A}_n$$

estuviera relacionado con el conjunto original por la ley de transformación (5). Este énfasis explícito en el conjunto o sistema de funciones y en la ley de transformación es lo que caracteriza el enfoque de Ricci de la teoría de invariantes diferenciales. El conjunto de funciones A_j, que son, por cierto, las componentes del gradiente de la función escalar A, constituye un ejemplo de tensor covariante de rango 1. La idea de un conjunto o sistema de funciones que caracterizan una cantidad invariante no era en sí nada nuevo, porque los vectores eran ya bien conocidos en tiempos de Ricci. Efectivamente, los vectores vienen representados por sus componentes en un sistema de coordenadas, y también están sujetos a una ley de transformación del mismo tipo, si es que el vector ha de permanecer, como es debido, invariante bajo un cambio de coordenadas. Sin embargo, los nuevos sistemas que introducía Ricci eran mucho más

generales, y el énfasis puesto en la ley de transformación también resultaba nuevo.

Como otro ejemplo del punto de vista introducido por Ricci, consideremos la expresión del elemento de longitud, que viene dada por

$$ds^2 = \sum_{i,j=1}^{n} g_{ij} dx^i dx^j. \tag{6}$$

Bajo un cambio de coordenadas, el valor de la distancia ds debe permanecer invariante, por razones puramente geométricas. Sin embargo, si realizamos la transformación (2) y escribimos la nueva expresión en la forma

$$\overline{ds^2} = \sum_{i,j=1}^{n} G_{ij} dy^i dy^j, \tag{7}$$

entonces $g_{ij}(x^1, x^2, \ldots, x^n)$ no será igual, en general, a $G_{ij}(y^1, y^2, \ldots, y^n)$ (cuando los valores de las y^i representen el mismo punto que las correspondientes x^i). Lo que sí se verifica es que

$$G_{kl} = \sum_{i,j=1}^{n} g_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} \tag{8}$$

donde en las funciones g_{ij} se han sustituido las x^i por sus valores en términos de las y^i. Para ver que efectivamente se verifica (8) no tenemos más que sustituir dx^i en (6) por

$$dx^i = \sum_{k=1}^{n} \frac{\partial x^i}{\partial y^k} dy^k$$

y dx^j por

$$dx^j = \sum_{l=1}^{n} \frac{\partial x^j}{\partial y^l} dy^l$$

y obtener el coeficiente de $dy^k dy^l$. Así pues, aunque G_{kl} no tiene por qué coincidir con g_{kl}, sabemos cómo obtener las G_{kl} a partir de las

g_{kl}. El conjunto de n^2 coeficientes g_{ik} de la forma cuadrática fundamental constituye otro tensor, un tensor covariante de rango 2, cuya ley de transformación viene dada por (8).

Ricci introdujo también los tensores contravariantes. Consideremos la transformación inversa de la (2) que hemos manejado hasta ahora; si esa inversa es

$$y^j = g_j(x^1, x^2, ..., x^n), \tag{9}$$

entonces

$$dy^j = \sum_{k=1}^{n} \frac{\partial y^j}{\partial x^k} dx^k. \tag{10}$$

Si ahora consideramos a las dx^k como un conjunto de cantidades que constituyen un tensor, entonces vemos que, desde luego, $dy^j \neq dx^j$, pero podemos obtener las dy^j a partir de las dx^j mediante la ley de transformación (10). El conjunto de elementos dx^k recibe por tanto el nombre de tensor contravariante de rango 1, donde el adjetivo «contravariante» indica la presencia en la transformación de las derivadas parciales $\partial y^j/\partial x^k$ en lugar de las derivadas $\partial x^i/\partial y^j$ que aparecían en (5) y (8). Así pues, las diferenciales mismas de las variables que se transforman constituyen un tensor contravariante de rango 1.

También podemos tener un tensor de rango 2 que se transforme de manera contravariante en ambos índices. Si el conjunto de funciones $A^{kl}(x^1,x^2, ..., x^n)$, $k,l = 1, 2, ..., n$ se transforma al realizar el cambio de variables (9) de manera que

$$\overline{A}^{ij} = \sum_{k,l=1}^{n} \frac{\partial y^i}{\partial x^k} \frac{\partial y^j}{\partial x^l} A^{kl}, \tag{11}$$

entonces ese conjunto constituye un tensor contravariante de rango 2. Además, podemos tener lo que se llaman tensores mixtos, que se transforman de manera covariante en algunos índices y de manera contravariante en los demás. Por ejemplo, el conjunto de funciones A^k_{ij}, $i, j, k = 1, 2, ..., n$, denota un tensor mixto en el que, de acuerdo con Ricci, los índices inferiores son aquellos en los que se transforma covariantemente, y el índice superior en el que se transforma contravariantemente. El tensor cuyos elementos son las A^k_{ij} se dice

que es de rango 3. Podemos tener tensores covariantes, contravariantes y mixtos de rango r. Un tensor n-dimensional de rango r tendrá n^r componentes. La ecuación (21) del capítulo 37 muestra que el símbolo de cuatro índices de Riemann (rk,ih) es un tensor covariante de rango 4. Un tensor covariante de rango 1 es un vector. Levi-Civita le asociaba el vector contravariante que se define como sigue: si el conjunto de las λ_i son las componentes de un vector covariante, entonces el conjunto de las

$$\lambda^i = \sum_{k=1}^{n} g^{ik}\lambda_k$$

constituye el vector contravariante asociado a él; g^{ik} es un cociente cuyo numerador es el menor complementario de g_{ik} en el determinante de las g_{ik} y cuyo denominador es g, el valor del determinante.

Se pueden realizar operaciones con los tensores. Por ejemplo, si tenemos dos tensores del mismo tipo, es decir, que tengan el mismo número de índices covariantes y contravariantes, podemos sumarlos, sumando las componentes con idénticos índices. Así,

$$A_i^j + B_i^j = C_i^j.$$

Se debe y puede mostrar que las C_i constituyen un tensor covariante en el índice i y contravariante en el índice j.

También se pueden multiplicar dos tensores cualesquiera cuyos índices vayan de 1 al mismo n. Un ejemplo bastará para mostrar cómo funciona esa multiplicación. Si hacemos

$$A_i^h B_j^k = C_{ij}^{hk}$$

se puede comprobar que el tensor con n^4 componentes C_{ij} es covariante en los índices inferiores y contravariante en los superiores. No hay una operación de división para los tensores.

Lo que sí puede realizarse con ellos es una operación llamada «contracción», que ilustraremos con el siguiente ejemplo: dado el tensor de componentes A_{ir}, definimos las cantidades

$$B_i^h = \sum_{r=1}^{n} A_{ir}^{hr},$$

y puede probarse entonces que el conjunto de cantidades B_i constituye un tensor de rango 2, covariante en el índice i, y contravariante en el índice h.

Resumiendo, un tensor es un conjunto de funciones (componentes) fijas con respecto a un sistema de referencia, o de coordenadas, que se transforman al realizar un cambio de éstas de acuerdo con ciertas leyes. Cada componente en un sistema de coordenadas es una función lineal homogénea de las componentes en otro sistema de coordenadas. Si las componentes de un tensor son iguales a las de otro en un determinado sistema de referencia, también son iguales en cualquier otro sistema de referencia. En particular, si las componentes se anulan en algún sistema de referencia, se anulan en cualquier otro. La igualdad de tensores es por tanto un invariante con respecto al cambio de coordenadas. El significado físico, geométrico, o incluso puramente matemático, que tiene un tensor en un sistema de referencia determinado, es preservado por la transformación, manteniéndolo en cualquier otro sistema de referencia. Esta propiedad es vital en la teoría de la relatividad, en la que cada observador posee su propio sistema de referencia; como las verdaderas leyes físicas son aquellas que se cumplen para todos los observadores, para reflejar esa independencia del sistema de coordenadas, esas leyes se expresan como tensores.

Disponiendo del concepto de tensor se puede reexpresar en forma tensorial muchos de los conceptos de la geometría riemanniana. El más importante es quizá la curvatura del espacio. La curvatura riemanniana (vid. cap. 37, sec. 3) se puede expresar tensorialmente de muchas maneras. Las expresiones modernas utilizan el convenio de sumación introducido por Einstein, según el cual se sobreentiende la suma siempre que aparece un índice repetido en el producto de dos símbolos. Así, por ejemplo,

$$g^{ij}\lambda_j = \sum_{j=1}^{n} g^{ij}\lambda$$

Con esta notación, el tensor de curvatura (vid. [20] del cap. 37) es:

$$R_{\lambda\mu\varrho\sigma} = \frac{\partial}{\partial x^{\varrho}}[\mu\sigma, \lambda] - \frac{\partial}{\partial x^{\sigma}}[\mu\varrho, \lambda] + \{\mu\varrho, \varepsilon\}[\lambda\sigma, \varepsilon] - \{\mu\sigma, \varepsilon\}[\lambda\varrho, \varepsilon]$$

o lo que es equivalente,

$$R^i_{jkl} = \frac{\partial}{\partial x^l}\{jk, i\} - \frac{\partial}{\partial x^k}\{jl, i\} - [\{sk, i\}\{jl, s\} - \{sl, i\}\{jk, s\}]$$

donde los corchetes denotan símbolos de Christoffel de primera especie, y las llaves símbolos de Christoffel de segunda especie. Cualquiera de esas formas es lo que se llama ahora tensor de curvatura de Riemann-Christoffel. A causa de ciertas relaciones (que no nos entretendremos en describir ahora) entre las componentes, el número de componentes distintas de ese tensor es $n^2(n^2 - 1)/12$. Para $n = 4$, que es el caso en la teoría general de la relatividad, ese número de componentes distintas es 20. En una variedad riemanniana bidimensional todas las componentes son iguales, por ejemplo, a R_{1212}. En este caso la curvatura total de Gauss K no es sino

$$K = \frac{R_{1212}}{g}$$

donde g es el determinante de las g_{ij}, es decir, $g_{11}g_{22} - g_{12}^2$. Si todas las componentes son nulas, el espacio es euclídeo.

A partir del tensor de Riemann-Christoffel, Ricci obtuvo por contracción el que se llama ahora tensor de Ricci o tensor de Einstein. Sus componentes R_{jl} son $\Sigma_{k=1}^n R_{jlk}$. Este tensor fue utilizado por Einstein,[3] con $n = 4$, para expresar la curvatura de su geometría riemanniana espacio-temporal.

3. Derivación covariante

Ricci también introdujo en el análisis tensorial una operación a la que él y (más tarde) Levi-Civita llamaron derivación covariante.[4] Esta operación había aparecido ya en los trabajos de Christoffel y Lipschitz;[5] Christoffel había presentado un método (vid. cap. 37, sec. 4) con el que se podían obtener, a partir de invariantes diferenciales formados con las derivadas de la forma fundamental para ds^2 y de ciertas funciones $\gamma(x_1, x_2, ..., x_n)$, invariantes asociados a deri-

---

[3] Zeit. für Math. und Phys., 62, 1914, 225-261.

[4] Atti della Accad. dei Lincei, Rendiconti, (4), 3, 1887, 15-18 = *Opere*, 1, 199-203.

[5] Jour. für Math., 70, 1869, 46-70 y 241-245, y 71-102.

vadas de orden superior. Ricci constantó la importancia de ese método para su análisis tensorial, y lo adoptó.

Mientras que Christoffel y Lipschitz realizaban la derivación covariante de la forma completa, Ricci, que ponía más énfasis en las componentes de los tensores, trabajó sobre ellas. En particular, si $A_1(x_1,x_2, ..., x_n)$ es una componente covariante de un vector o tensor de rango 1, la derivada covariante de A_i no es simplemente su derivada con respecto a x^l, sino el tensor de rango 2

$$A_{i,l} = \frac{\partial A_i}{\partial x_l} - \sum_{j=1}^{n} \{il, j\} A_j, \tag{12}$$

donde las llaves indican el símbolo de Christoffel de segunda especie. De igual modo, si A_{ik} es una componente de un tensor covariante de rango 2, su derivada covariante con respecto a x^l viene dada por

$$A_{ik,l} = \frac{\partial A_{ik}}{\partial x^l} \quad \sum_{j=1}^{n} \{il, j\} A_{jk} - \sum_{j=1}^{n} \{kl, j\} A_{ij} \ , \tag{13}$$

Para un tensor contravariante de rango 1 con componentes A^i, la derivada covariante $A^i_{,l}$ viene dada por

$$A^i_{,l} = \frac{\partial A^i}{\partial x^l} + \sum_{j=1}^{n} A^j\{jl, i\},$$

y es un tensor mixto de segundo orden. Para el tensor mixto con componentes A^h_i la derivada covariante es

$$A^h_{i,l} = \frac{\partial A^h_i}{\partial x^l} - \sum_{j=1}^{n} A^h_j\{il, j\} + \sum_{j=1}^{n} A^j_i\{jl, h\}.$$

La derivada covariante de un invariante escalar γ es el vector covariante cuyas componentes vienen dadas por $\gamma_i = \partial_\gamma/\partial x_i$. A ese vector se le llama gradiente del invariante escalar.

Desde el punto de vista puramente matemático, la derivada covariante de un tensor es otro tensor cuyo rango es una unidad mayor en los índices covariantes. Se trata de un hecho importante, ya que

posibilita el tratamiento de tales derivadas en el marco general del análisis tensorial. También tiene significado geométrico: supongamos que tenemos un campo vectorial constante en el plano, esto es, un conjunto de vectores, anclado cada uno de ellos en un punto distinto, pero con la misma magnitud y dirección. En tal caso, las componentes con respecto a un sistema rectangular de coordenadas son también constantes. Sin embargo, las componentes de esos vectores con respecto al sistema polar de coordenadas, es decir, una componente a lo largo del radio vector y la otra perpendicular a ese radio vector cambian de un punto a otro, porque las direcciones en que se toman esas componentes también cambian de un punto a otro. Si se obtienen las derivadas con respecto a las coordenadas, digamos r y ϑ, de esas componentes, el cambio expresado por esas derivadas, que ya no son nulas, refleja el cambio en las componentes debido al sistema de coordenadas, y no un cambio en los vectores mismos. Las coordenadas utilizadas en geometría riemanniana son curvilíneas; el efecto de la curvatura de esas coordenadas viene dado por los símbolos de Christoffel de segunda especie (denotados aquí mediante llaves). La derivada covariante completa de un tensor representa la tasa de cambio efectiva de la cantidad física o geométrica representada por ese tensor, así como el cambio debido al sistema de coordenadas.

En los espacios euclídeos, en los que la ds^2 se puede siempre reducir a una suma de cuadrados con coeficientes constantes, la derivada covariante se reduce a la ordinaria, ya que los símbolos de Christoffel son nulos. También es nula la derivada covariante de cada g_{ij} que aparece en la métrica riemanniana. Esto fue demostrado por el mismo Ricci,[6] y se conoce como lema de Ricci.

El concepto de derivación covariante nos permite expresar con facilidad para los tensores generalizaciones de nociones ya conocidas en análisis vectorial, que pueden tratarse ahora en geometría riemanniana. Así, si $A_i(x_1, x_2, \ldots, x_n)$ son las componentes de un vector n-dimensional A, entonces

$$\theta = \sum_{i,l=1}^{n} g^{il} A_{i,l}, \tag{14}$$

[6] Atti della Accad. dei Lincei, Rendiconti, (4), 5, 1889, 112-118 = *Opere*, 1, 268-275.

donde g^{il} ha sido ya definido más arriba, es un invariante diferencial. Cuando la métrica fundamental es la de un sistema rectangular de coordenadas en un espacio euclídeo (coordenadas cartesianas), las constantes g^{il} son nulas excepto para $i = l$, y en este caso las derivadas ordinaria y covariante son idénticas, (14) se convierte entonces en

$$\theta = \sum_{i=1}^{n} \frac{\partial A_i}{\partial x^i}$$

y constituye el análogo n-dimensional de lo que en tres dimensiones se llama la divergencia. Por eso (14) se llama también la divergencia del tensor cuyas componentes son las A_i. También se puede mostrar, utilizando (14), que si A es una función escalar, la divergencia del gradiente de A viene dada por

$$\Delta_2 A = \frac{1}{\sqrt{g}} \sum_{l=1}^{n} \frac{\partial}{\partial x^l} (\sqrt{g}\, A^l), \tag{15}$$

donde

$$A^l = \sum_{i=1}^{n} g^{il} \frac{\partial A}{\partial x^i} = \sum_{i=1}^{n} g^{il} A_i.$$

Esta es también la expresión de Beltrami para $\Delta_2 A$ en una geometría riemanniana (vid. cap. 37, sec. 5).

Aunque Ricci y Levi-Civita dedicaron una gran parte de su artículo de 1901 a la técnica del análisis tensorial, su principal objetivo era la búsqueda de invariantes diferenciales. Plantearon el siguiente problema general: dada una forma diferencial cuadrática positiva γ y un número arbitrario de funciones asociadas S, determinar todos los invariantes diferenciales absolutos que se pueden formar a partir de los coeficientes de γ, las funciones S, y las derivadas de esos coeficientes y funciones hasta un orden dado m; ofrecieron una solución completa: basta encontrar los invariantes *algebraicos* del sistema formado por la forma diferencial cuadrática fundamental γ, las derivadas covariantes de las funciones asociadas S, hasta el orden m y, si $m > 1$, una cierta forma tetralineal G_4 cuyos coeficientes son los símbolos de Riemann (ih,jk) y sus derivadas covariantes hasta el orden $m - 2$.

Concluían su artículo mostrando cómo podían expresarse ciertas ecuaciones en derivadas parciales y leyes físicas en forma tensorial, haciéndolas así independientes del sistema de coordenadas. Este era el objetivo declarado por Ricci. El análisis tensorial se utilizó así para expresar la invarianza matemática de leyes físicas, muchos años antes de que Einstein lo empleara con el mismo propósito.

4. El desplazamiento paralelo

Desde 1901 hasta 1905 la investigación en análisis tensorial se vio limitada a un muy pequeño grupo de matemáticos. No obstante, la obra de Einstein cambió el panorama. Albert Einstein (1879-1955), cuando aún trabajaba como ingeniero en la oficina suiza de patentes, provocó una gran excitación en el mundo científico con el anuncio de su teoría de la relatividad especial, o restringida.[7] En 1914 aceptó un puesto en la Academia Prusiana de Ciencias en Berlín, como sucesor del célebre químico-físico Jacobus Van't Hoff (1852-1911). Dos años más tarde anunció su teoría general de la relatividad.[8]

Las propuestas revolucionarias de Einstein acerca de la relatividad de los fenómenos físicos despertaron un intenso interés entre los físicos, filósofos y matemáticos de todo el mundo. Los matemáticos se sintieron especialmente atraídos por la naturaleza de la geometría que Einstein había encontrado conveniente utilizar para la formulación de sus teorías.

La exposición de la teoría restringida, que implica las propiedades de variedades pseudoeuclídeas tetradimensionales (espacio-tiempo), se hace más cómoda con la ayuda de vectores y tensores; pero la teoría general, que depende de las propiedades de las variedades riemannianas tetradimensionales, exige el uso del cálculo tensorial especial asociado a tales variedades. Afortunadamente ese cálculo ya se había desarrollado, aunque no había atraído hasta entonces la atención de los físicos.

El trabajo de Einstein sobre la teoría restringida no utiliza de hecho geometría riemanniana ni análisis tensorial.[9] Pero esa teoría

---

[7] *Annalen der Phys.*, 17, 1905, 891-921; hay traducción inglesa en la edición de Dover de A. Einstein, *The principle of relativity* (1951).

[8] Annalen der Phys., 49, 1916, 769-822.

[9] La métrica es $ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2$. Se trata de un espacio de curvatura constante. Toda sección por un plano t = const. es euclídea.

restringida no tenía en cuenta la acción de la gravitación. Einstein comenzó entonces a trabajar sobre el problema de la fuerza gravitatoria y de cómo dar cuenta de su efecto imponiendo una estructura en la geometría espacio-temporal que hiciera que los objetos se movieran automáticamente a lo largo de los mismos caminos que los que se derivarían de tener en cuenta la acción de la fuerza gravitacional. En 1911 hizo pública una teoría que interpretaba de esa manera el efecto de una fuerza gravitatoria que tuviera dirección constante en todo el espacio, sabiendo desde luego que esa teoría no era realista. Hasta ese momento Einstein había utilizado únicamente los instrumentos matemáticos más simples, objetando incluso la necesidad de la «matemática elevada», de la que sospechaba que a menudo se introducía sólo para dificultar la lectura. Sin embargo, tratando de avanzar en su problema discutió sobre él en Praga con un colega, el matemático Georg Pick, que atrajo su atención hacia la teoría matemática de Ricci y Levi-Civita. En Zürich encontró a un amigo, Marcel Grossmann (1878-1936), que le ayudó a comprender esa teoría y, con ella como base, consiguió formular la teoría general de la relatividad.

Para representar su mundo tetradimensional de tres coordenadas espaciales y una cuarta representando el tiempo, Einstein utilizó la métrica riemanniana

$$ds^2 = \sum_{i,j=1}^{4} g_{ij}\, dx_i\, dx_j, \tag{16}$$

donde x_4 denota la coordenada temporal. Las g_{ij} debían elegirse de manera que reflejaran la presencia de materia en las diferentes regiones del universo. Además, como la teoría se ocupa de la determinación de longitudes, tiempos, masas y otras cantidades físicas, por diferentes observadores que se mueven de manera arbitraria unos con respecto a otros, los «puntos» del espacio-tiempo vienen representados en diferentes sistemas de coordenadas, ligado cada uno a un observador. La relación entre un sistema de coordenadas y otro viene dada por una transformación

$$x_i = \phi_i(y_1, y_2, y_3, y_4), \qquad i = 1, \ldots, 4.$$

Las leyes de la naturaleza son aquellas relaciones o expresiones que son iguales para todos los observadores. Así pues, son invariantes en el sentido matemático.

Desde el punto de vista de las matemáticas, la importancia de la obra de Einstein consistió, como ya hemos indicado, en la extensión del interés por el análisis tensorial y la geometría riemanniana. La primera innovación en análisis tensorial posterior a la teoría de la relatividad se debió a Levi-Civita. En 1917, mejorando una idea de Ricci, introdujo [10] el concepto de desplazamiento paralelo de un vector. Gerhard Hessenberg introdujo independientemente esa noción en el mismo año.[11] Brouwer la había empleado ya en 1906 para superficies de curvatura constante. El objetivo de este concepto consiste en definir lo que se entiende por vectores paralelos en una variedad riemanniana. Se puede constatar la dificultad de hacerlo considerando la superficie de una esfera, que con la distancia determinada por arcos de círculo máximo es una variedad riemanniana. Si un vector, partiendo digamos de un paralelo y apuntando hacia el norte (el vector debe ser tangente a la superficie esférica) se mueve a lo largo de esa circunferencia no máxima, manteniéndose paralelo a sí mismo en el espacio euclídeo tridimensional, cuando haya recorrido media circunferencia ya no será tangente a la esfera, y no pertenecerá al espacio en cuestión. Para obtener una noción de paralelismo de vectores válida en variedades riemannianas, hay que generalizar el concepto de euclídeo, aunque en el proceso se pierdan algunas de las propiedades usuales.

La idea geométrica utilizada por Levi-Civita para definir el transporte o desplazamiento paralelo se comprenderá más fácilmente para una superficie. Consideremos una curva C sobre la superficie, y desplacemos un vector anclado en un punto de C paralelamente a sí mismo en el siguiente sentido: en cada punto de C hay un plano tangente a la superficie. La envolvente de esa familia de planos es una superficie desarrollable, y cuando ésta se desarrolla sobre un plano, los vectores paralelos a lo largo de C deben ser paralelos en el plano euclídeo.

Levi-Civita generalizó esta idea a las variedades riemannianas n-dimensionales. Cuando un vector se desplaza paralelamente a sí mismo en el plano euclídeo a lo largo de una línea recta —una geodésica del plano— el vector forma siempre el mismo ángulo con la recta. De acuerdo con esto, el paralelismo en una variedad riemanniana se define así: cuando un vector se mueve a lo largo de una

[10] Rendiconti del Circolo Matematico di Palermo, 42, 1917, 173-205.

[11] Math. Ann., 78, 1918, 187-217.

geodésica, debe seguir formando el mismo ángulo con la geodésica (con la tangente a la geodésica). En particular, una tangente a la geodésica se mantiene paralela a sí misma cuando se desplaza a lo largo de ésta. Por definición, el vector debe mantener la misma magnitud. Se sobreentiende que el vector se mantiene tangente a la variedad riemanniana, incluso si ésta está inmersa en un espacio euclídeo. La definición de transporte paralelo requiere también que se mantenga el ángulo formado por dos vectores cuando ambos se desplazan paralelamente a lo largo de la misma curva C. En el caso general de transporte paralelo a lo largo de una curva cerrada arbitraria C, el vector inicial y el final no tendrán usualmente la misma dirección (euclídea). La desviación en la dirección dependerá del camino C. Consideremos por ejemplo un vector que parte de un punto P perteneciente a un paralelo C de una superficie esférica, tangente al correspondiente meridiano que pasa por P; cuando se realiza el *transporte paralelo* de ese vector a lo largo de C, al volver a P seguirá siendo tangente a la superficie esférica, pero formará un ángulo de valor $2\pi(1 - \cos\gamma)$ con el vector de partida, donde γ es la colatitud de P.

Si se utiliza la definición general de desplazamiento paralelo a lo largo de una curva de una variedad riemanniana, se obtiene una condición analítica. La ecuación diferencial que satisfacen las componentes X^α de un vector contravariante que se desplaza paralelamente a lo largo de una curva es (sobreentendiendo la suma en los índices repetidos):

$$\frac{dX^\alpha}{dt} + \{\beta\gamma, \alpha\} X^\beta \frac{du^\gamma}{dt} = 0, \qquad \alpha = 1, 2, \ldots, n, \tag{17}$$

donde las $u^i(t)$, $i = 1, 2, \ldots, n$ definen la curva. Para un vector covariante X_α la condición es

$$\frac{dX_\alpha}{dt} - \{\alpha l, j\} X_j \frac{du^l}{dt} = 0, \qquad \alpha = 1, 2, \ldots, n. \tag{18}$$

Esas ecuaciones se pueden utilizar para definir el transporte paralelo a lo largo de cualquier curva C. La solución determinada unívocamente por los valores de las componentes en un punto dado P es un vector con valores en cada punto de C y paralelo por definición

al vector inicial en P. La ecuación (18) establece que la derivada covariante de X_α es 0.

Una vez introducida la noción de desplazamiento paralelo, se puede describir en términos de él la curvatura del espacio; concretamente, en términos del cambio experimentado por un vector infinitesimal al ser transportado paralelamente a distancias infinitesimales. El paralelismo está en la base del concepto de curvatura incluso en un espacio euclídeo, ya que la curvatura de un arco infinitesimal depende del cambio de dirección de un vector tangente a lo largo del arco.

5. Generalizaciones de la geometría riemanniana

El éxito que tuvo la utilización de la geometría riemanniana en la teoría de la relatividad resucitó el interés por ella. Sin embargo, la obra de Einstein planteó una cuestión incluso más general: había incorporado el efecto gravitacional de la masa en el espacio utilizando las funciones adecuadas para las g_{ij}; como consecuencia, las geodésicas de su espacio-tiempo eran precisamente los caminos recorridos por objetos que se movieran libremente, como por ejemplo lo hace la Tierra en torno al Sol. A diferencia de lo que ocurría en la mecánica newtoniana, no se requería una fuerza gravitatoria para explicar el camino recorrido. La eliminación de la gravedad sugirió la posibilidad de explicar también la atracción y repulsión entre las cargas eléctricas en términos de la métrica del espacio. Tal resultado proporcionaría una teoría unificada de la gravitación y el electromagnetismo. Estos trabajos condujeron a generalizaciones de la geometría riemanniana conocidas colectivamente con el nombre de geometrías no-riemannianas.

En geometría riemanniana la ds^2 liga entre sí los diferentes puntos del espacio, especificando la relación entre ellos por la distancia que los separa. En las geometrías no-riemannianas la conexión entre diferentes puntos se especifica de otras maneras, que no tienen que basarse necesariamente en una métrica. La variedad de tales geometrías es muy amplia, y cada una de ellas tiene un desarrollo tan extenso como la misma geometría riemanniana. Aquí daremos sólo algunos ejemplos de las ideas básicas de esas geometrías.

Quien inició primeramente este área de trabajo fue Hermann

Weyl,[12] y las geometrías que introdujo se conocen como espacios con una conexión afín. En geometría riemanniana, la demostración de que la derivada covariante de un tensor es también un tensor depende sólo de relaciones de la forma

$$\{\overline{ik, h}\} = \{ab, j\} \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^k} \frac{\partial y^h}{\partial x^j} + \frac{\partial^2 x^j}{\partial y^i \partial y^k} \frac{\partial y^h}{\partial x^j}, \qquad (19)$$

en donde el miembro de la izquierda es la transformada de $\{ab,j\}$ por el cambio de coordenadas que pasa de las x^i a las y^l. Los símbolos de Christoffel, que se definen en términos de los coeficientes de la forma fundamental, satisfacen esas relaciones. Consideremos en su lugar funciones L^i_{jk} y $\overline{L}^i_{jk}$ de las x^i e y^i, respectivamente, que satisfagan las mismas relaciones (19) pero sin estar necesariamente relacionadas con la forma cuadrática fundamental. Un conjunto de funciones L^i_{jk} asociadas a un espacio V_n que tengan la propiedad de transformación (19) constituyen lo que se llama una conexión afín; las funciones mismas se llaman coeficientes de la conexión afín, y se dice que el espacio V_n posee una conexión afín, o que es un espacio afín. La geometría riemanniana resulta como caso especial cuando los coeficientes de la conexión afín son los símbolos de Christoffel de segunda especie, que se obtienen a partir del tensor fundamental del espacio. Dadas las funciones L, es posible introducir conceptos tales como derivación covariante, curvatura, y otras nociones análogas a las de la geometría riemanniana. Sin embargo, en esa nueva geometría ya no se puede hablar de la magnitud de un vector.

Es un espacio con conexión afín, una curva tal que sus tangentes sean paralelas con respecto a la curva (en el sentido del desplazamiento paralelo en ese espacio) es lo que se llama un camino propio del espacio. Esos caminos propios constituyen una generalización de las geodésicas de una variedad riemanniana. Todos los espacios conectados afínmente que tienen las mismas funciones L tienen los mismos caminos propios. La geometría de los espacios conectados afínmente puede prescindir así de la métrica riemanniana. Weyl obtuvo las ecuaciones de Maxwell a partir de las propiedades del espacio, pero la teoría en su conjunto no casaba demasiado bien con otros hechos establecidos.

Otra geometría no riemanniana, debida a Luther P. Eisenhart

[12] Mathematische Zeitschrift, 2, 1918, 384-411 = *Ges. Abh.*, 2, 1-28.

(1876-1965) y Veblen,[13] llamada geometría de los caminos, funciona de manera un poco diferente: se parte de n^3 funciones $\Gamma^i_{\lambda\mu}$ de x^1, ..., x^n; el sistema de n ecuaciones diferenciales

$$\frac{d^2x^i}{ds^2} + \sum_{\lambda\mu} \Gamma^i_{\lambda\mu} \frac{dx^\lambda}{ds} \frac{dx^\mu}{ds} = 0, \qquad i = 1, 2, ..., n, \tag{20}$$

con $\Gamma^i_{\lambda\mu} = \Gamma^i_{\mu\lambda}$ define entonces una familia de curvas, llamadas caminos, que son las geodésicas de esa geometría (en la geometría riemanniana las ecuaciones (20) son precisamente las de las geodésicas); a partir de ellas se puede entonces construir la geometría de los caminos de una manera análoga a la de la geometría riemanniana.

Otra generalización distinta de la geometría riemanniana se debe a Paul Finsler (1894), quien la expuso en su excelente tesis de 1918 en Götingen.[14] La ds^2 riemanniana se sustituye por una función más general $F(x,dx)$ de las coordenadas y sus diferenciales, sobre la que se imponen restricciones que aseguren la posibilidad de minimizar la integral $\int F(x,(dx/dt))dt$ obteniendo así las geodésicas.

No obstante, todos los intentos realizados hasta el momento por generalizar el concepto de geometría riemanniana, incorporando tanto el electromagnetismo como los fenómenos gravitatorios, han resultado infructuosos. Aun así, los matemáticos siguen trabajando en esas geometrías abstractas.

Bibliografía

Cartan, E.: «Les récentes généralisations de la notion d'espace», Bull. des Sci. Math., 48, 1924, 294-320.

Pierpont, James: «Some modern views of space», Amer. Math. Soc. Bull., 32, 1926, 225-258.

Ricci-Curbastro, G.: *Opere*, 2 vol., Edizioni Cremonese, 1956-1957.

Ricci-Curbastro, G. y T. Levi-Civita: «Méthodes de calcul différential absolu et leurs applications», Math. Ann., 54, 1901, 125-201.

Thomas, T. Y.: «Recent trends in geometry», *Amer. Math. Soc. Semicentennial Publications*, II, 1938, 98-135.

Weatherburn, C. E.: *The development of multidimensional differential geo-*

[13] Proceedings of the National Academy of Sciences, 8, 1922, 19-23.

[14] Esta tesis se publicó hace algún tiempo: *Uber Kurven und Flachen in allgemeinen Raumen*, Birkhauser Verlag, Basilea, 1951.

metry, Australian and New Zealand Association for the Advancement of Science, 21, 1933, 12-28.

Weitzenbock, R.: «Neuere arbeiten der algebraischen Invariantentheorie. Differentialinvarianten», *Encyk. der Math. Wiss.*, B. G. Teubner, 1902-1927, II, parte III, E1, 1-71.

Weyl, H.: *Mathematische Analyse des Raumproblems* (1923), Chelsea (reimpresión), 1964.

Capítulo 49
LA APARICION DEL ALGEBRA ABSTRACTA

> Quizá pueda, sin inmodestia, reclamar para mí mismo la denominación de Adán Matemático, porque creo que he dado más nombres (que han pasado a ser de uso general) a criaturas de la razón matemática que todos los restantes matemáticos de la época juntos.
>
> J. J. SYLVESTER

1. El panorama del siglo XIX

En el campo del álgebra abstracta, lo mismo que en la mayoría de los desarrollos matemáticos del siglo XX, tanto los conceptos como las metas fundamentales se fijaron a lo largo del siglo XIX. El hecho básico de que el álgebra puede tratar sistemas de objetos que no sean necesariamente números reales o complejos, quedó definitivamente demostrado en una buena docena de creaciones del siglo XIX. Ejemplos de objetos que se combinaban mediante operaciones y bajo leyes para dichas operaciones peculiares del sistema en cuestión son los vectores, cuaterniones, matrices, formas tales como la $ax^2 + bxy + cy^2$, hipernúmeros de diversos tipos, transformaciones, y sustituciones o permutaciones. Incluso en la teoría de números algebraicos, a pesar de que trata de clases de números complejos, aparecían en escena diversidad de álgebras, porque estas clases sólo tenían algunas de las propiedades del sistema completo de los números complejos.

Estas diversas clases de objetos se distinguieron y se clasificaron de acuerdo con las propiedades de las operaciones definidas sobre ellas, y ya hemos visto que se introdujeron nociones tales como las

de grupo, anillo, ideal y cuerpo, y nociones subordinadas como las de subgrupo, subgrupo invariante o extensión de un cuerpo, para identificar conjuntos concretos de propiedades. Sin embargo, a lo largo de casi todo el siglo XIX los trabajos sobre esos diversos tipos de álgebras se referían a los sistemas concretos que hemos mencionado más arriba. Solamente durante las últimas décadas del siglo se dieron cuenta los matemáticos de que podían ascender a un nivel de eficacia nuevo integrando juntas muchas álgebras hasta entonces dispersas, por abstracción de su contenido común. Así, por ejemplo, los grupos de permutaciones, los grupos de clases de formas estudiados por Gauss, los hipernúmeros con su suma y los grupos de transformaciones podían ser estudiados todos de un solo golpe considerando un conjunto de cosas o elementos sometidos a una operación cuya naturaleza quedaba especificada únicamente por ciertas propiedades abstractas, siendo la primera y más importante de las cuales que dicha operación, aplicada a dos elementos del conjunto, produce otro elemento del mismo conjunto. Y las mismas ventajas podían conseguirse para las diversas colecciones que constituían anillos o cuerpos, evidentemente. Aunque la idea de trabajar con colecciones abstractas fue anterior a las axiomáticas de Pasch, Peano y Hilbert, los desarrollos posteriores en este sentido aceleraron indudablemente la aceptación general de los planteamientos abstractos dentro del álgebra.

Así surgió el álgebra abstracta como el estudio explícitamente consciente de clases enteras de álgebras, las cuales, individualmente consideradas, no sólo eran sistemas concretos sino que servían para fines también concretos en áreas específicas de la matemática, como ocurría con los grupos de sustituciones en la teoría de ecuaciones. La ventaja de obtener resultados que podían ser útiles en muchas ramas concretas de la matemática considerando versiones abstractas, pronto se perdió de vista, y el estudio de las estructuras abstractas y de sus propiedades se convirtió en un fin en sí mismo.

El álgebra abstracta ha sido uno de los campos más favorecidos de la matemática del siglo XX, y hoy sus dominios son muy extensos. Nosotros sólo presentaremos aquí los comienzos de esta materia e indicaremos las posibilidades casi ilimitadas que en ella ofrece la investigación. La mayor dificultad para exponer lo que se ha ido haciendo en este campo es la de la terminología. Aparte de las dificultades usuales de que diferentes autores usen términos distintos y de que los términos cambien de significado de un período a otro, el

álgebra abstracta viene marcada por la introducción de cientos de términos nuevos. Toda variación en un concepto, por pequeña que sea, aparece distinguida por un término nuevo e impresionante; un diccionario completo de los términos utilizados llenaría un libro de buen tamaño.

2. La teoría abstracta de grupos

La primera estructura abstracta que se introdujo y estudió fue la de grupo. Muchas de las ideas básicas de la teoría de grupos abstractos se pueden encontrar ya, tanto implícita como explícitamente, en época tan temprana como el 1800, por lo menos. Una de las actividades favoritas de los historiadores de la matemática, ahora que existe la teoría abstracta, es la de rastrear cuántas de las ideas abstractas aparecen ya anunciadas en las obras concretas de Gauss, Abel, Galois, Cauchy y docenas de otros. Aquí no dedicaremos espacio a esta reconstrucción del pasado. El único punto importante que vale la pena mencionar es el de que, una vez adquirida la noción abstracta, fue relativamente fácil para los fundadores de la teoría de grupos abstractos obtener ideas y resultados reformulando obras del pasado.

Antes de examinar el desarrollo del concepto abstracto de grupo, puede ser interesante saber qué es lo que pretendían los matemáticos. La definición abstracta de grupo que se suele utilizar hoy se refiere a una colección de elementos, en cantidad finita o infinita, y a una operación que, aplicada a dos elementos cualesquiera de la colección, produce otro de la misma (propiedad de cierre). La operación es asociativa; hay un elemento tal que para cualquier elemento a del grupo $ae = ea = a$; y para cada elemento a existe un elemento inverso a' tal que $a'a = aa' = e$. Cuando la operación es conmutativa, el grupo se llama conmutativo o abeliano, y la operación suele llamarse suma y representarse por $+$; el elemento e se representa entonces por 0 y se le llama elemento cero. Si la operación no es conmutativa se le suele llamar multiplicación, y el elemento e se representa por 1 y se llama identidad.

El concepto de grupo abstracto y las propiedades que lo caracterizan surgieron lentamente. Podemos recordar (cap. 31, sec. 6) que Cayley había propuesto ya la idea de grupo abstracto en 1849, pero

en esta época no se reconoció su importancia. En 1858 Dedekind,[1] muy por delante de su época, dio una definición abstracta de los grupos finitos, derivada de los grupos de permutaciones. En 1877[2] observó de nuevo que sus módulos de números algebraicos, a los que pertenecían $\alpha + \beta$ y $\alpha - \beta$ si pertenecían α y β, se podían generalizar de manera que los elementos no fueran ya números algebraicos, y la operación podía ser arbitraria con tal de tener un inverso y ser conmutativa; así pues, sugería un grupo finito abstracto conmutativo. La visión de Dedekind del valor de la abstracción es notable. Se dio cuenta claramente, trabajando en la teoría de números algebraicos, del valor de estructuras tales como los ideales y los cuerpos; él fue el verdadero creador del álgebra abstracta.

Kronecker,[3] siguiendo los trabajos sobre los números ideales de Kummer, dio también lo que podemos considerar la definición abstracta de un grupo abeliano finito, análoga a la de Cayley de 1849. Kronecker habla expresamente de elementos abstractos, una operación abstracta, la propiedad de clausura, las propiedades asociativa y conmutativa y la existencia de un inverso único para cada elemento, y a continuación demuestra varios teoremas. Entre las diversas potencias de un elemento cualquiera θ, hay una que es igual al elemento unidad 1; si ν es el mínimo exponente para el que $\theta^{\nu} = 1$, entonces para cada divisor μ de ν hay un elemento ϕ tal que $\phi^{\mu} = 1$. Si ϕ^{ϱ} y ϕ^{σ} son ambos iguales a 1 y ϱ y σ son los mínimos exponentes para lo que esto ocurre y son primos entre sí, entonces $(\theta\phi)^{\varrho\sigma} = 1$. Knonecker también dio la primera demostración de lo que hoy se llama un teorema de la base: existe un sistema finito fundamental de elementos θ_1, θ_2, θ_3, ... tales que los productos

$$\theta_1^{h_1}\theta_2^{h_2}\theta_3^{h_3}\ldots, \qquad h_i = 1, 2, 3, \ldots, n_i$$

representan todos los elementos del grupo exactamente una vez. Los mínimos valores posibles de n_1, n_2, n_3 ... que corresponden a θ_1, θ_2, θ_3, ... (es decir, para los cuales $\theta^{i} = 1$) son tales que cada uno es divisible por el siguiente y el producto $n_1\, n_2\, n_3 \ldots$ es igual al número de elementos n del grupo. Además todos los factores primos de n están en n_1.

[1] *Werke*, 3, 439-446.
[2] Bull. des Sci. Math., (2), 1, 1877, 17-41, en particular p. 41 = *Werke*, 3, 262-296.
[3] Monatsber. Berliner Akad., 1870, 881-889 = *Werke*, 1, 271-282.

En 1878 escribió Cayley cuatro artículos más sobre grupos abstractos finitos.[4] En ellos, como en los de 1849 y 1854, subraya que un grupo puede ser considerado como un concepto general y no necesita limitarse a los grupos de sustituciones aunque, señala, todo grupo (finito) pueda ser representado como un grupo de sustituciones. Estos artículos de Cayley tuvieron más influencia que los anteriores, porque la época estaba ya madura para una abstracción que abarcaba más que los grupos de sustituciones.

En un artículo conjunto de Frobenius y Ludwig Stickelberger (1850-1936) [5] se da el importante paso de reconocer que el concepto abstracto de grupo incluye las congruencias y la composición de formas de Gauss, así como los grupos de sustituciones de Galois. Se menciona además la existencia de grupos de orden infinito.

Eugen Netto, en su libro *Substitutionentheorie und ihre Anwendung auf die Algebra* (1882) se limitaba a tratar grupos de sustituciones, pero los enunciados de sus conceptos y teoremas permitían reconocer el carácter abstracto de dichos conceptos. Aparte de reunir resultados de sus predecesores, Netto trata de los conceptos de isomorfismo y homomorfismo. El primero significa una correspondencia biunívoca entre dos grupos, tal que si $ab = c$, donde a, b y c son elementos del primer grupo, entonces $a'b' = c'$, siendo a', b' y c' los elementos correspondientes del segundo grupo. Un homomorfismo es una correspondencia en general, tal que de nuevo $ab = c$ implica $a'b' = c'$.

Hacia 1880 surgieron nuevas ideas sobre grupos. Klein, influenciado por la obra de Jordan sobre grupos de permutaciones, había demostrado en su Erlanger Programm (cap. 38, sec. 5) que se podía usar grupos infinitos de transformaciones para clasificar las geometrías. Estos grupos son además continuos en el sentido de que en cualquier grupo se incluyen transformaciones arbitrariamente pequeñas o, dicho de otra manera, los parámetros de las transformaciones pueden tomar todos los valores reales. Así, en las transformaciones que expresan giros de los ejes, el ángulo θ puede tomar todos los valores reales. En sus trabajos sobre funciones automorfas, Klein y

---

[4] Math. Ann., 13, 1878, 561-565; Proc. London Math. Soc., 9, 1878, 126-133; Amer. J. of Math., 1, 1878, 50-52 y 174-176: todo el volumen 10 de sus *Collected Math. Papers*.

[5] Jour. für Math., 86, 1879, 217-262 = Frobenius, *Ges. Abh.*, 1, 545-590.

Poincaré habían utilizado otro grupo de tipo infinito, un grupo no continuo o discreto (cap. 29, sec. 6).

Sophus Lie, que había trabajado con Klein en torno a 1870, adoptó el concepto de grupo continuo de transformaciones, pero para otro objetivo que el de clasificar las geometrías. Había observado que la mayor parte de las ecuaciones diferenciales ordinarias que se habían integrado por viejos métodos eran invariantes bajo clases de grupos de transformaciones continuas, y pensó que esto podría arrojar luz sobre la resolución de ecuaciones diferenciales y clasificarlas.

En 1874 introdujo Lie su teoría general de grupos de transformaciones.[6] Un grupo viene representado por un sistema de ecuaciones

$$x_i' = f_i(x_1, x_2, ..., x_n, a_1, ..., a_n), \qquad i = 1, 2, ..., n, \tag{1}$$

donde las f_i son funciones analíticas de las x_i y a_i. Las a_i son parámetros mientras que las x_i son las variables, y $(x_1, x_2, ..., x_n)$ representa un punto del espacio n-dimensional. Tanto los parámetros como las variables pueden tomar todos los valores reales o complejos. Por ejemplo, en una dimensión, la clase de transformaciones

$$x' = \frac{ax + b}{cx + d}$$

donde a, b, c y d pueden tomar todos los valores reales, es un grupo continuo. Los grupos representados por (1) se llaman finitos, donde la palabra finito se refiere al número de parámetros; el número de transformaciones es infinito, desde luego. El caso unidimensional anterior es un grupo tri-paramétrico porque sólo importan las razones de a, b y c a d. En el caso general, el producto de dos transformaciones

$$x_i' = f_i(x_1, ..., x_n, a_1, ..., a_n)$$
$$x_i'' = f_i(x_1', ..., x_n', b_1, ..., b_n)$$

es la transformación

$$x_i'' = f_i(x_1, ..., x_n, c_1, ..., c_n)$$

[6] Nachrichten König. Ges. der Wiss. zu Gött., 1874, 529-542 = *Ges. Abh.*, 5, 1-8.

donde las c_i son funciones de las a_i y las b_i. En el caso de una variable, Lie llamó al grupo una variedad simplemente extendida, y en el de n variables una variedad múltiplemente extendida.

En un artículo de 1883 sobre grupos continuos, publicado en una obscura revista noruega,[7] introduce Lie los grupos continuos de transformaciones infinitos. Estos no vienen definidos por ecuaciones tales como la (1), sino por medio de ecuaciones diferenciales. Las transformaciones resultantes no dependen ahora de un número finito de parámetros continuos, sino de funciones arbitrarias. No hay un concepto abstracto de grupo correspondiente a estos grupos continuos infinitos, y aunque se ha investigado mucho sobre ellos, no los consideraremos aquí.

Quizá sea de interés observar que, al comienzo de sus trabajos, Klein y Lie definían un grupo de transformaciones como uno que posee sólo la propiedad de clausura. Las otras propiedades, tales como la existencia de inversa de cada transformación, elemento neutro, etc. se establecían utilizando las propiedades de las transformaciones o, como en el caso de la propiedad asociativa, se utilizaban como propiedades obvias de las transformaciones. Lie reconoció durante el curso de su trabajo que habría que postular como parte de la definición de un grupo la existencia de inverso de cada elemento.

Hacia 1880 se conocían ya cuatro tipos principales de grupos. Los grupos discontinuos de orden finito, como los grupos de sustituciones; los grupos discontinuos (o discretos) infinitos, tales como los que aparecen en la teoría de funciones automorfas; los grupos continuos finitos de Lie, como los grupos de transformaciones de Klein y las transformaciones analíticas más generales de Lie; y los grupos continuos infinitos de Lie definidos por ecuaciones diferenciales.

Con la obra de Walther von Dyck (1856-1934), las tres raíces principales de la teoría de grupos, la teoría de ecuaciones, la teoría de números y los grupos de transformaciones infinitos, quedaron todas incluidas en el concepto abstracto de grupo. Dyck estaba influenciado por Cayley y fue discípulo de Felix Klein. En 1882 y 1883 [8] publicó artículos sobre teoría de grupos abstractos, que incluía los grupos continuos y discretos. Su definición de grupo se refiere a un conjunto de elementos y una operación que satisfaga la

---

[7] *Ges. Abh.*, 5, 314-360.

[8] Math. Ann. 20, 1882, 1-44 y 22, 1883, 70-118.

propiedad de clausura, la asociativa, y la existencia de inverso de cada elemento, pero no la conmutativa.

Dyck estudia de manera explícita el concepto de generadores de un grupo, que estaba implícito ya en el teorema de la base de Kronecker y explícito en los trabajos de Netto sobre grupos de sustituciones. Los generadores constituyen un subconjunto fijo de elementos independientes de un grupo, tales que todo elemento del mismo se puede expresar como producto de potencias de los generadores y sus inversos. Cuando no hay restricción alguna sobre los generadores el grupo se llama un grupo libre. Si los generadores son A_1, A_2, ... entonces una expresión de la forma

$$A_1^{\mu_1} A_2^{\mu_2} \ldots,$$

donde los μ_i son enteros positivos o negativos, se llama una palabra. Puede haber relaciones entre los generadores, que serán de la forma

$$F_i(A_i) = 1;$$

es decir, una palabra o combinación de palabras igual al elemento unidad del grupo. Dyck demuestra entonces que la presencia de relaciones implica un subgrupo invariante y un grupo cociente $\bar{G}$ de grupo libre G. En su artículo de 1883 aplica la teoría de grupos abstractos a los grupos de permutaciones, grupos finitos de rotaciones (simetrías de poliedros), grupos de la teoría de números y grupos de transformaciones.

Huntington,[9] E. H. Moore,[10] y Leonard E. Dickson (1874-1954) [11] dieron todos ellos conjuntos de postulados independientes para el concepto de grupo abstracto; éstos, así como otros sistemas de postulados, son en realidad pequeñas variaciones unos de otros.

Habiendo conquistado ya la noción abstracta de grupo, los matemáticos se dedicaron a demostrar teoremas sobre grupos abstractos sugeridos por resultados conocidos para casos concretos. Así, Frobenius [12] demostró el teorema de Sylow (cap. 31, sec. 6) para grupos

---

[9] Amer. Math. Soc. Bull., 8, 1902, 296-300 y 388-391, y Amer. Math. Soc. Trans. 6, 1905, 181-197.

[10] Amer. Math. Soc. Trans., 3, 1902, 485-492, y 6, 1905, 179-180.

[11] Amer. Math. Soc. Trans., 6, 1905, 198-204.

[12] Jour für Math., 100, 1887, 179-181 = *Ges. Abh.*, 2, 301-303.

abstractos finitos: todo grupo finito cuyo orden, es decir, su número de elementos, sea divisible por la potencia ν de un primo p, contiene siempre un subgrupo de orden p^ν.

Aparte de buscar grupos concretos para propiedades que podían verificarse en grupos abstractos, muchos matemáticos introdujeron conceptos directamente para los grupos abstractos. Dedekind [13] y George A. Miller (1863-1951) [14] estudiaron los grupos no abelianos en que todo subgrupo es normal (o invariante). Dedekind, en su artículo de 1897 y Miller [15] introdujeron el concepto de conmutador y subgrupo conmutador; si s y t son elementos de un grupo G, al elemento $s^{-1}t^{-1}st$ se le llama conmutador de s y t. Tanto Dedekind como Miller utilizaron esta concepto en sus teoremas; por ejemplo, el conjunto de todos los conmutadores de los pares ordenados de elementos de un grupo G generan un subgrupo invariante de G. Hölder [16] y E. H. Moore [17] estudiaron de manera abstracta los automorfismos de un grupo, es decir, las transformaciones biunívocas de un grupo en sí mismo bajo las cuales si $ab = c$ entonces $a'b' = c'$.

El desarrollo posterior de la teoría de grupos abstractos ha seguido direcciones muy distintas. Una de ellas partía de los grupos de sustituciones en el artículo de Hölder de 1893 y pretendía hallar todos los grupos de un orden dado, problema que ya había mencionado Cayley en sus artículos de 1878.[18] El problema general se ha resistido a ser resuelto y, en consecuencia se han investigado órdenes particulares, tales como p^2q^2 donde p y q son primos. Un problema relacionado con éste ha sido el de la enumeración de los grupos intransitivos, primitivos e imprimitivos de diversos grados (el número de letras en un grupo de sustituciones).

Otra dirección de investigación ha sido la determinación de grupos compuestos o solubles y de los grupos simples, es decir, aquellos que no tienen subgrupos invariantes aparte de la identidad. Este problema tiene su origen en la teoría de Galois, naturalmente. Hölder, después de introducir el concepto de grupo cociente,[19] estudió

---

[13] Math. Ann., 48, 1897, 548-561 = *Werke*, 2, 87-102.
[14] Amer. Mat. Soc. Bull., 4, 1898, 510-515 = *Coll. Works*, 1, 266-269.
[15] Amer. Math. Soc. Bull., 4, 1898, 135-139 = *Coll. Works*, 1, 254-257.
[16] Math. Ann., 43, 1893, 301-412.
[17] Æmer. Math. Soc. Bull., 1, 1895, 61-66, y 2, 1896, 33-43.
[18] *Coll. Math. Papers*, 10, 403.
[19] Math. Ann., 34, 1889, 26-56.

los grupos simples [20] y compuestos.[21] Entre sus resultados está el de que un grupo cíclico de orden primo es simple, y así le ocurre al grupo alternado de todas las permutaciones pares de n letras para $n \geqslant 5$. Se han encontrado muchos otros grupos finitos simples.

En cuanto a los grupos resolubles, Frobenius dedicó varios artículos al problema. Descubrió, por ejemplo,[22] que todos los grupos cuyo orden no sea divisible por el cuadrado de un primo son resolubles.[23] El problema general de investigar qué grupos son resolubles es parte del problema más general de determinar la estructura de un grupo dado.

En sus artículos de 1882 y 1883 había introducido Dyck la idea abstracta de grupo definido por generadores y relaciones entre ellos. Dado un grupo definido en términos de un número finito de generadores y relaciones, el problema de la identidad o de las palabras, formulado por Max Dehn,[22] es el de determinar si una «palabra» o producto de elementos cualesquiera es igual al elemento unidad. Puede darse cualquier conjunto de relaciones, porque, en el peor de los casos, el grupo trivial consistente sólo en la identidad las satisface. Decidir si un grupo dado por generadores y relaciones es trivial, no es trivial; de hecho, no hay ningún procedimiento efectivo para hacerlo. Para una relación demostró Wilhelm Magnus (1907) que el problema es resoluble,[25] pero el problema general no lo es.[26]

Otro famoso problema sin resolver de la teoría general de grupos es el problema de Burnside. Todo grupo finito tiene las propiedades de ser finitamente generado y de que todo elemento tiene orden finito; en 1902 [27] William Burnside (1852-1927) se preguntó si el recíproco sería cierto, es decir, si un grupo G es finitamente generado y si todo elemento tiene orden finito, ¿es G necesariamente

---

[20] Math. Ann., 40, 1892, 55-88 y 43, 1893, 301-412.

[21] Math. Ann., 46, 1895, 321-422.

[22] Sitzungsber. Akad. Wiss. zu Berlin, 1893, 337-345, y 1895, 1027-1044 = *Ges. Abh.*, 2, 565-573, 677-694.

[23] Un resultado muy importante ha sido obtenido recientemente por Walter Feit (1930) y John G. Thompson (Pacific Jour. of Math., 13, parte 2, 1963, 775-1029): todos los grupos finitos de orden impar son resolubles. Burnside ya había sugerido en 1906 que podía ser así.

[24] Math. Ann., 71, 1911, 116-144.

[25] Math. Ann., 106, 1932, 295-307.

[26] Esto fue demostrado por P. S. Novikov en 1955. Véase American Math. Soc. Translations (2), 9, 1958, 1-122.

[27] Quat. Jour. of Math., 33, 230-238.

finito? Este problema ha atraído mucha atención, pero solamente se han resuelto casos particulares. Otro problema afín, el problema del isomorfismo, consiste en determinar cuándo dos grupos, definido cada uno por generadores y relaciones, son isomorfos.

Uno de los giros sorprendentes de la teoría de grupos ha sido el de que, poco después de romper amarras la teoría abstracta, los matemáticos se dedicaron a buscar representaciones por medio de álgebras más concretas para obtener resultados relativos a los grupos abstractos. Cayley había señalado ya en su artículo de 1854 que todo grupo abstracto finito puede ser representado por un grupo de permutaciones. También hemos mencionado (cap. 31, sec. 6) que Jordan había introducido en 1878 la representación de grupos de sustituciones por transformaciones lineales. Estas transformaciones o sus matrices han demostrado ser la representación más eficaz de los grupos abstractos y reciben el nombre de representaciones lineales.

Una representación matricial de un grupo G es un homorfismo de los elementos g de G en un conjunto de matrices cuadradas no singulares $A(g)$ de orden fijo y de elementos complejos. El homomorfismo implica que

$$A(g_i g_j) = A(g_i)A(g_j)$$

para todo g_i y g_j de G. Hay muchas representaciones matriciales posibles de un mismo grupo G, alterando el orden (o dimensión) de las matrices, e incluso para un orden dado puede variar la correspondencia establecida. También se pueden sumar dos representaciones; si para cada elemento g de G es A_g la matriz correspondiente en una representación de orden m y B_g la matriz correspondiente en otra de orden n, entonces

$$\begin{pmatrix} A_g & 0 \\ 0 & B_g \end{pmatrix}$$

es otra representación llamada suma de las representaciones por separado. Análogamente, si

$$E_g = \begin{pmatrix} B_g & C_g \\ 0 & D_g \end{pmatrix}$$

es otra representación, cuando B_g y D_g son matrices no singulares

de órdenes m y n respectivamente, las B_g y D_g también son representaciones y de orden menor que la E_g. A E_g se le llama una representación graduada; tanto ella como cualquier representación equivalente a ella (F_g^{-1} E_g F_g es una representación equivalente si F_g es no singular y del mismo orden que E_g) se llama reducible. Una representación no equivalente a otra graduada se llama irreducible. La idea básica de una representación irreducible consistente en un conjunto de transformaciones lineales en n variables, es que es una representación homomorfa o isomorfa en la que es imposible elegir $m < n$ funciones lineales de las variables que se transformen entre sí bajo toda operación del grupo que representan. Una representación equivalente a una suma de representaciones irreducibles se llama completamente reducible.

Todo grupo finito tiene una representación especial llamada regular. Supongamos que los elementos del grupo son $g_1, g_2, ..., g_n$. Sea a uno cualquiera de los g y consideremos una matriz $n \times n$; supongamos que a $g_i = g_j$. Entonces colocamos un 1 en el lugar (i,j) de la matriz, para todo g_i y el elemento fijo a, añadiendo ceros en todos los lugares restantes de la matriz. La matriz así obtenida corresponde al elemento a. Esta matriz está definida para cada g del grupo y el conjunto constituye una representación regular por la izquierda; análogamente, formando los productos $g_i a$ obtendríamos una representación regular por la derecha, y reordenando los elementos g del grupo resultarían otras representaciones regulares. La idea de una representación regular fue introducida por Charles S. Peirce en 1879.[28]

La representación de grupos de sustituciones por transformaciones lineales de la forma

$$x_i' = \sum_j a_{ij}x_j, \qquad i = 1, 2, ..., n$$

iniciada por Jordan, fue generalizada al estudio de representaciones de todos los grupos finitos por Frobenius, Burnside, Theodor Molien (1861-1941) e Issai Schur (1875-1941) a finales del siglo XIX y comienzos del XX. Frobenius[29] introdujo el concepto de representación reducible y completamente reducible para grupos finitos, y

---

[28] Amer. J. of Math., 4, 1881, 221-225.

[29] Sitzungsber, Akad. Wiss. zu Berlin, 1897, 994-1015 = *Ges. Abh.*, 3, 82-103.

demostró que una representación regular contiene todas las representaciones irreducibles. En otros artículos publicados entre 1897 y 1910, algunos conjuntamente con Schur, demostró muchos otros resultados, incluyendo el hecho de que hay sólo unas pocas representaciones irreducibles, con las que se componen todas las demás.

Burnside [30] obtuvo otro resultado importante, una condición necesaria y suficiente sobre los coeficientes de un grupo de transformaciones lineales en *n* variables para que el grupo sea reducible. El hecho de que cualquier grupo finito de transformaciones lineales es completamente reducible lo demostró por primera vez Heinrich Maschke (1853-1908).[31] La teoría de representaciones de grupos finitos ha conducido a importantes resultados sobre grupos abstractos. Durante el segundo cuarto de este siglo la teoría de representación se generalizó a grupos continuos, pero aquí no entraremos en ello.

En el estudio de las representaciones de grupos es importante el concepto de carácter de un grupo, que puede hacerse remontar a la obra de Gauss, Dirichlet y Heinrich Weber (véase nota 35). Esta noción la formuló de manera abstracta Dedekind para grupos abelianos en la tercera edición de las *Vorlesungen über Zahlentheorie* (1879) de Dirichlet. Un carácter de un grupo es una función $x(s)$ definida para todos los elementos s, tal que no es cero para ningún s y $x(ss') = x(s)\, x(s')$. Dos caracteres son distintos si $x(s) \neq x'(s)$ para un s al menos del grupo.

Esta definición fue generalizada para todos los grupos finitos por Frobenius. Después de enunciar una definición bastante complicada,[32] dio otra más sencilla,[33] que hoy es la habitual. La función carácter es la traza (o suma de los elementos de la diagonal principal) de las matrices de una representación irreducible del grupo. Frobenius mismo y otros aplicarían más tarde el mismo concepto a grupos infinitos.

Los caracteres de los grupos suministran entre otras cosas una determinación del número mínimo de variables en términos de las cuales se puede representar un grupo finito como un grupo de trans-

---

[30] Proc. London Math. Soc., (2), 3, 1905, 430-434.

[31] Math. Ann., 52, 1899, 363-368.

[32] Sitzungsber. Akad. Wiss. zu Berlin, 1896, 985-1021 = *Ges. Abh.*, 3, 1-37.

[33] Sitzungsber. Akad. Wiss. zu Berlin, 1897, 994-1015 = Ges. Abh., 3, 82-103.

formaciones lineales, y para grupos conmutativos permiten determinar todos los subgrupos.

Muchos matemáticos de finales del siglo XIX y comienzos del XX mostraron el entusiasmo usual por las corrientes de moda, pensando que toda la matemática que valía la pena recordar terminaría incluida en la teoría de grupos. Klein en particular, a pesar de que no le gustaba el formalismo de la teoría *abstracta* de grupos, se mostró favorable al concepto de grupo porque pensaba que unificaría la matemática. Poincaré se mostró igualmente entusiasta, diciendo [34] «... la teoría de grupos es, por así decirlo, la totalidad de la matemática desprovista de su materia y reducida a pura forma».

3. La teoría abstracta de cuerpos

El concepto del cuerpo R generado por n cantidades a_1, a_2, ..., a_n, es decir, el conjunto de todas las cantidades formadas sumando, restando, multiplicando y dividiendo repetidamente estas cantidades (exceptuada la división por cero), así como el concepto del cuerpo extensión obtenido añadiendo un nuevo elemento λ no en R, aparecen ya en la obra de Galois. Sus cuerpos eran los dominios de racionalidad de los coeficientes de una ecuación, y las extensiones se construían añadiendo una raíz. El mismo concepto tuvo también un origen muy distinto en los trabajos sobre números algebraicos de Dedekind y Kronecker (cap. 34, sec. 3) y, de hecho, el nombre «cuerpo» *(Körper)* se debe a Dedekind.

La teoría abstracta de cuerpos la inició Heinrich Weber, que ya había adoptado el punto de vista abstracto en teoría de grupos. En 1893 [35] dio una exposición abstracta de la teoría de Galois en la que introducía cuerpos (conmutativos) como extensiones de grupos. Un cuerpo, dice Weber, es una colección de elementos sometida a dos operaciones, llamadas suma y multiplicación, que satisfacen la condición de clausura, las propiedades asociativa y conmutativa, y la distributiva; además, cada elemento debe tener un inverso único para cada operación, excepto la división por cero. Weber insiste en que los conceptos de grupo y de cuerpo son los dos más importantes

[34] Acta Math., 38, 1921, 145.

[35] Math. Ann., 43, 1893,521-549. Para los grupos, ver Math. Ann., 20, 1882, 301-329.

del álgebra. Algo más tarde Dickson [36] y Huntington [37] dieron sistemas de axiomas independientes para un cuerpo.

A los cuerpos que ya se conocían en el siglo XIX, los de los números racionales, reales y complejos, los cuerpos de números algebraicos y los cuerpos de funciones racionales en una o varias variables, añadió Kurt Hensel otro tipo, los cuerpos p-ádicos, que abrían nuevos campos en teoría de números algebraicos (*Theorie der algebraischen Zahlen*, 1908). Lo primero que observó Hensel fue que cualquier entero D se puede expresar de una y sólo una manera como suma de potencias de un número primo p. Es decir

$$D = d_0 + d_1p + ... + d_kp^k$$

donde d_i es algún entero de 0 a $p - 1$. Por ejemplo,

$$14 = 2 + 3 + 3^2$$
$$216 = 2 \cdot 3^3 + 2 \cdot 3^4.$$

Análogamente, cualquier número racional r (distinto de cero) puede escribirse en la forma

$$r = \frac{a}{b} p^n$$

donde a y b son enteros no divisibles por p y n es un entero positivo, negativo o nulo. Hensel generalizó estas observaciones e introdujo los números p-ádicos, que son expresiones de la forma

$$\sum_{i=-\rho}^{\infty} c_i p^i \tag{2}$$

donde p es un número primo y los coeficientes c_i son números racionales irreducibles cuyo denominador no es divisible por p. Tales expresiones no tendrán en general valor numérico ordinario; sin embargo, son entes matemáticos por definición.

Hensel define las cuatro operaciones básicas con estos números y demuestra que forman un cuerpo. Un subconjunto de los números

[36] Amer. Math. Soc. Trans., 4, 1903, 13-20 y 6, 1905, 198-204.
[37] Amer. Math. Soc. Trans., 4, 1903, 31-37 y 6, 1905, 181-197.

p-ádicos puede ponerse en correspondencia biunívoca con los números racionales y, de hecho, este subconjunto es isomorfo a los números racionales en el sentido de isomorfismo entre cuerpos. Hensel define en el cuerpo de los números p-ádicos las unidades, los números enteros p-ádicos y otros conceptos análogos a los de los números racionales ordinarios.

Definiendo los polinomios cuyos coeficientes sean números p-ádicos, puede hablar Hensel de raíces p-ádicas de ecuaciones polinómicas y extender a estas raíces todos los conceptos de los cuerpos de números algebraicos. Así, habrá enteros algebraicos p-ádicos y, más en general, números algebraicos p-ádicos, y pueden formarse cuerpos de números algebraicos p-ádicos que sean extensiones de los números «racionales» p-ádicos definidos por (2). De hecho, toda la teoría ordinaria de números algebraicos puede trasladarse a los números p-ádicos. De una manera un tanto sorprendente, la teoría de números algebraicos p-ádicos conduce a resultados sobre los números algebraicos ordinarios, y también resulta útil al estudiar las formas cuadráticas y ha conducido al concepto de cuerpo valuado.

La variedad creciente de cuerpos indujo a Ernst Steinitz (1871-1928), que estaba muy influenciado por la obra de Weber, a emprender un estudio sistemático de los cuerpos abstractos, cosa que hizo en su artículo fundamental *Algebraischen Theorie der Körper.*[38] Todos los cuerpos, según Steinitz, pueden dividirse en dos tipos principales. Sea K un cuerpo y consideremos todos los subcuerpos de K (por ejemplo, los números racionales son un subcuerpo de los números reales). Los elementos comunes a todos los subcuerpos constituyen un subcuerpo llamado el subcuerpo primo P de K. Hay dos tipos posibles de cuerpos primos; el elemento unidad e está contenido en P y, por tanto, lo están

$$e, 2e, \dots, ne, \dots$$

Estos elementos, o bien son todos distintos o existe un entero p tal que $pe = 0$. En el primer caso P tiene que contener todas las fracciones ne/me y, dado que estos elementos forman un cuerpo P ha de ser isomorfo al cuerpo de los números racionales, y se dice que K tiene característica 0.

Si en cambio se tiene $pe = 0$, es fácil demostrar que el más pe-

[38] Jour. für Math., 137, 1910, 167-309.

queño p que cumple esta condición tiene que ser primo, y el cuerpo es isomorfo al de los restos de los enteros módulo p, es decir 0, 1, 2, ..., $p-1$. Entonces se dice que K es un cuerpo de característica p y cualquier subcuerpo suyo tiene la misma característica. En este caso $pa = pea = 0$, es decir, todas las expresiones en K pueden reducirse módulo p.

A partir del cuerpo primo P, en cualquiera de los dos casos anteriores, se puede obtener el cuerpo original K por un proceso de adjunción de elementos. El método consiste en tomar un elemento a de K que no esté en P y formar todas las funciones racionales $R(a)$ de a con coeficientes en P y después, si es necesario, tomar un b que no esté en $R(a)$ y hacer lo mismo con b, y continuar el proceso repetidamente.

Si uno parte de un cuerpo arbitrario K, pueden hacerse diversos tipos de adjunciones. Una adjunción simple se obtiene añadiendo un único elemento x. El cuerpo extendido tiene que contener todas las expresiones de la forma

$$a_0 + a_1x + ... + a_nx^n \tag{3}$$

donde los a_i son elementos de K. Si estas expresiones son todas distintas, el cuerpo extendido es el cuerpo $K(x)$ de todas las funciones racionales en x con coeficientes en K. Tal adjunción se llama una adjunción transcendente, y $K(x)$ una extensión transcendente de K. Si algunas de las expresiones (3) son iguales, se puede demostrar que existe una relación (utilizando α en lugar de x)

$$f(\alpha) = \alpha^m + b_1\alpha^{m-1} + ... + b_m = 0$$

con las b_i en K y $f(x)$ irreducible sobre K. Entonces las expresiones

$$C_1\alpha^{m-1} + ... + C_m$$

con los C_i en K constituyen un cuerpo $K(\alpha)$ formado por la adjunción de α a K. Este cuerpo se llama entonces una extensión algebraica simple de K. En $K(\alpha)$, $f(x)$ tiene una raíz, y recíprocamente, si tomamos un polinomio irreducible arbitrario $f(x)$ sobre K, entonces se puede construir un $K(\alpha)$ en el que $f(x)$ tenga una raíz.

Un teorema fundamental de Steinitz afirma que todo cuerpo se puede obtener a partir de su cuerpo primo haciendo primero una

serie de adjunciones transcendentes (posiblemente infinitas), y a continuación una serie de adjunciones algebraicas al cuerpo transcendente. Un cuerpo K' se llama extensión algebraica de K si se obtiene por sucesivas adjunciones algebraicas simples, y si su número es finito se dice que K' es de rango finito.

No todo cuerpo se puede extender por adjunciones algebraicas. Por ejemplo, en el caso de los complejos es imposible porque todo polinomio $f(x)$ es reducible sobre este cuerpo; un cuerpo con esta propiedad se llama algebraicamente cerrado. Steinitz demostró también que para todo cuerpo K existe un único cuerpo algebraicamente cerrado K' que es algebraico sobre K en el sentido de que cualquier otro cuerpo algebraicamente cerrado sobre K (que contenga a K) contiene un subcuerpo isomorfo a K'.

Steinitz estudió también el problema de determinar en qué cuerpos se verifica la teoría de Galois. Decir que la teoría de Galois se verifica en un cuerpo significa lo siguiente: un cuerpo de Galois $\bar{K}$ sobre un cuerpo dado K es un cuerpo algebraico en el que todo polinomio irreducible $f(x)$ en K, o bien permanece irreducible, o se descompone en un producto de factores lineales. Para todo cuerpo de Galois $\bar{K}$ existe un conjunto de automorfismos, cada uno de los cuales transforma los elementos de $\bar{K}$ en otros del mismo cuerpo y tales que a $\alpha \pm \beta$ y a $\alpha\beta$ corresponden $\alpha' \pm \beta'$ y $\alpha'\beta'$ mientras que todos los elementos de K quedan invariantes (se corresponden a sí mismos). Este conjunto de automorfismos forma un grupo G, llamado grupo de Galois de $\bar{K}$ con respecto a K. El teorema principal de la teoría de Galois afirma que hay una única correspondencia entre los subgrupos de G y los subcuerpos de $\bar{K}$ tal que a cada subgrupo G' de G corresponde el subcuerpo K' de todos los elementos invariantes por G' y recíprocamente. Se dice que la teoría de Galois se verifica para aquellos cuerpos que cumplen este teorema. El resultado de Steinitz dice esencialmente que la teoría de Galois se verifica en los cuerpos de rango finito que pueden obtenerse de un cuerpo dado por una serie de adjunciones de raíces de polinomios irreducibles $f(x)$ que no tengan raíces iguales. Los cuerpos en los que todos los $f(x)$ irreducibles no tienen raíces iguales se llaman separables (Steinitz los llamó *volkommen* o completos).

La teoría de cuerpos incluye también, como indica la clasificación de Steinitz, cuerpos finitos de característica p. Un ejemplo sencillo de este tipo es el conjunto de todos los restos módulo un número primo p. El concepto de cuerpo finito se debe a Galois; en

1830 éste publicó un importante artículo, «Sur la théorie des nombres».[39] Galois pretendía resolver en él las congruencias

$$F(x) \equiv 0 \bmod p,$$

donde p es primo y $F(x)$ un polinomio de grado n. Galois tomó un $F(x)$ irreducible (módulo p), de manera que la congruencia no tenía raíces enteras ni irracionales. Esto le obligó a considerar otras soluciones, sugeridas por los números imaginarios. Representó una de las raíces de $F(x)$ por i (que no es $\sqrt{-1}$) y consideró la expresión

$$a_0 + a_1 i + a_2 i^2 + \ldots + a_{n-1} i^{n-1} \qquad (4)$$

donde los a_i son números enteros. Cuando estos coeficientes varían sobre los restos positivos módulo p, esta expresión toma solamente p^ν valores. Sea α uno de estos valores no nulos, de los que habrá $p^\nu - 1$. Las potencias de α también tienen la forma (4), luego no pueden ser todas distintas. Debe haber al menos una potencia $\alpha^m = 1$, con m mínimo; entonces habrá m valores distintos

$$1, \alpha, \alpha^2, \ldots, \alpha^{m-1}. \qquad (5)$$

Si multiplicamos estas m cantidades por una expresión β de la misma forma, obtenemos un grupo de cantidades distintas de las (5) y entre sí. Multiplicando el conjunto (5) por γ hallaremos otras cantidades, hasta que obtengamos todas las de la forma (4). Por tanto, m tiene que dividir a $p^\nu - 1$ o bien $\alpha^{p^\nu - 1} = 1$ y por tanto $\alpha^p = \alpha$. Los p^ν valores de la forma (4) constituyen un cuerpo finito. Así pues, Galois ha demostrado en esta situación concreta que el número de elementos de un cuerpo de Galois de característica p es una potencia de p.

E. H. Moore [40] demostró que cualquier cuerpo finito es isomorfo a un cuerpo de Galois de orden p^n, con p primo. Existe tal cuerpo para todo primo p y todo entero positivo n y su característica es p. Joseph H. M. Wedderburn (1882-1948), profesor de la universidad de Princeton,[41] y Dickson demostraron simultáneamente que todo

[39] Bulletin des Sciences Mathématiques de Férussac, 13, 1830, 428-435 = *Œuvres*, 1897, 15-23.

[40] N. Y. Math. Soc. Bull., 3, 1893, 73-78.

[41] Amer. Math. Soc. Trans., 6, 1905, 349-352.

cuerpo finito es conmutativo (para la multiplicación). Se ha investigado mucho tratando de determinar la estructura de los grupos aditivos de los cuerpos de Galois, y de los cuerpos mismos.

4. Anillos

Aunque las estructuras de anillo e ideal eran bien conocidas y las utilizaron Dedekind y Kronecker en sus trabajos sobre números algebraicos, la teoría abstracta es en su totalidad un producto del siglo XX. El nombre ideal ya se había utilizado (cap. 34, sec. 4), pero Kronecker utilizó la palabra «orden» para anillo, término este último introducido por Hilbert.

Antes de pasar a la historia, será bueno precisar el significado moderno de los conceptos. Un anillo abstracto es una colección de elementos que forma un grupo abeliano con respecto a una operación llamada suma y sobre los que está definida una segunda operación de multiplicación. Esta segunda operación es asociativa, pero puede ser conmutativa o no; puede haber o no un elemento unidad y además se verifica la propiedad distributiva $a(b + c) = ab + ac$ y $(b + c)a = ba + ca$.

Un ideal de un anillo R es un subanillo M tal que si a pertenece a M y r es un elemento cualquiera de R entonces ar y ra pertenecen a M. Si solamente ar pertenece a M, entonces M se llama un ideal por la derecha, y si sólo pertenece a M ra, un ideal por la izquierda. Si un ideal lo es por la derecha y por la izquierda se llama bilátero. El anillo total es un ideal llamado idea unidad, y el ideal (a) generado por un elemento a está formado por todos los elementos $ra + na$ donde r pertenece a R y n es un número entero. Si R tiene unidad, entonces $ra + na = ra + nea = (r + ne)$ a $= r'a$, donde r' es un elemento de R. A estos ideales generados por un elemento se les llama ideales principales. Cualquier ideal distinto de 0 y de R recibe el nombre de ideal propio. Análogamente al caso de un generador, si $a_1, a_2, ..., a_m$ son m elementos dados del anillo R con elemento unidad, entonces el conjunto de todas las sumas $r_1a_1 + r_2a_2 + ... + r_ma_m$ con coeficientes r_i en R constituyen un ideal de R, que se representa por $(a_1, a_2, ..., a_m)$. Se trata del ideal más pequeño que contiene a los elementos $a_1, a_2, ..., a_m$. Un anillo conmutativo R se llama noetheriano cuando todo ideal es de esta forma.

Cualquier ideal M en un anillo R, al ser un subgrupo del grupo

aditivo del anillo, divide a dicho anillo en clases de restos. Dos elementos a y b de R son congruentes con respecto a M, $a \equiv b$ (mod M), si $a - b$ pertenece a M. Si T es un homomorfismo del anillo R en el anillo R', lo que supone que $T(a + b) = Ta + Tb$, $T(ab) = Ta \cdot Tb$ y $T1 = 1'$, los elementos de R que se aplican en el elemento 0 de R' constituyen un ideal llamado núcleo de T, y R' es isomorfo al anillo cociente de R por el núcleo de T. Recíprocamente, dado un ideal L en R podemos formar el anillo cociente de R por L y definir un homomorfismo de R en R cociente L que tiene como núcleo L.

En la definición de anillo no se exige la existencia de inverso de cada elemento para la multiplicación. Si existe elemento unidad y también dicho inverso (excepto para 0), el anillo se llama un anillo con división o cuerpo no conmutativo *(skew field)*. Ya hemos mencionado que Wedderburn demostró en 1905 que todo anillo con división finito es un cuerpo conmutativo. Hasta 1905 las únicas álgebras con división conocidas eran los cuerpos conmutativos y los cuaterniones. Entonces Dickson introdujo otras nuevas, tanto conmutativas como no conmutativas. En 1914 el mismo Dickson [42] y Wedderburn [43] dieron los primeros ejemplos de cuerpos no conmutativos con centros (conjunto de todos los elementos que conmutan con todos los demás) de rango n^2.[44]

A finales del siglo XIX se introdujo una gran variedad de álgebras lineales asociativas concretas (cap. 32, sec. 6). Estas álgebras, consideradas en abstracto, son anillos, y cuando se formuló la teoría de anillos abstractos, ésta incluyó y generalizó los trabajos sobre estas álgebras concretas. Esta teoría de álgebras lineales asociativas, y en general todo el álgebra abstracta, recibió un nuevo impulso cuando Wedderburn, en su artículo «On Hypercomplex Numbers»,[45] generalizó resultados anteriores de Elie Cartan (1869-1951).[46] Recordemos que los números hipercomplejos son de la forma

---

[42] Amer. Math. Soc. Trans., 15, 1914, 31-46.

[43] Amer. Math. Soc. Trans., 15, 1914, 162-166.

[44] En 1958 demostraron Michel Kervaire (1927) (en los Proceedings of the National Academy of Sciences, 44, 1958, 280-283) y John Milnor (1931) (en Annals of Math., (2), 68, 1958, 444-449), utilizando ambos un resultado de Raoul Bott (1923), que las únicas álgebras con división posibles con coeficientes reales, si no se suponen las propiedades asociativa y conmutativa de la multiplicación, son los números reales, los complejos, los cuaterniones y los números de Cayley.

[45] Proc. London Math. Soc., (2), 6, 1907, 77-118.

[46] Ann. Fac. Sci. de Toulouse, 128, 1898, 1-99 = *Vuvres*, parte II, vol. I, 7-105

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n \tag{6}$$

donde las e_i son unidades formales y los x_i números reales o complejos. Wedderburn sustituyó los x_i por elementos de un cuerpo arbitrario F, y denominó a estas álgebras lineales asociativas generalizadas simplemente álgebras. Para estudiarlas tuvo que abandonar los métodos de sus predecesores debido a que el cuerpo arbitrario F no tenía por qué ser algebraicamente cerrado, aunque adoptó y perfeccionó la técnica de los idempotentes de Benjamin Peirce.

En el artículo de Wedderburn, pues, un álgebra consiste en todas las combinaciones lineales de la forma (6), ahora con coeficientes en un cuerpo F. El número de las e_i, llamadas unidades básicas, es finito y se denomina el orden del álgebra. La suma de dos de tales elementos viene dada por

$$\sum x_ie_i + \sum y_ie_i = \sum (x_i + y_i)e_i,$$

el producto escalar de un elemento a del cuerpo F y un elemento x del álgebra está definido por

$$a\sum x_ie_i = \sum ax_ie_i$$

y el producto de dos elementos del álgebra por

$$\left(\sum x_ie_i\right)\left(\sum y_je_j\right) = \sum_{ij} x_iy_je_ie_j,$$

complementado por una tabla que da todos los productos e_ie_j como combinaciones lineales de los e_k con coeficientes en F. Se exige que el producto sea asociativo, y siempre se puede añadir un elemento unidad (o módulo) 1 tal que $x \cdot 1 = 1 \cdot x = x$ para todo x, y entonces los elementos del álgebra incluyen los del cuerpo F de los coeficientes.

Dada un álgebra A, un subconjunto B de elementos que a su vez constituyan un álgebra, se llama una subálgebra. Si x pertenece a A e y a B y yx y xy pertenecen ambos a B, entonces B se llama una subálgebra invariante. Si el álgebra A es la suma de dos subálgebras invariantes sin elementos comunes, se llama reducible o también se dice que es la suma directa de las subálgebras.

Un álgebra simple es la que no tiene ninguna subálgebra invariante. Wedderburn también utilizó y modificó el concepto de álgebra semisimple de Cartan. Para definir este concepto hizo uso del elemento nilpotente: el elemento x es nilpotente si $x^n = 0$ para algún entero n, y se le llama propiamente nilpotente si tanto xy como yx son nilpotentes para todo y del álgebra A. Se puede demostrar que el conjunto de todos los elementos propiamente nilpotentes de un álgebra A forman una subálgebra invariante. Entonces, un álgebra semisimple A es la que no tiene ninguna subálgebra invariante nilpotente.

Wedderburn demostró que toda álgebra semisimple se puede expresar como suma directa de álgebras irreducibles, y cada una de ellas es equivalente al producto directo de un álgebra de matrices y un álgebra con división (o álgebra primitiva en la terminología de Wedderburn). Esto significa que cada elemento del álgebra irreducible puede considerarse como una matriz cuyos elementos pertenecen al álgebra con división. Dado que las álgebras semisimples se pueden reducir a una suma directa de varias álgebras simples, este teorema supone la determinación de todas las álgebras semisimples. Otro resultado utiliza el concepto de álgebra de matrices total, que es simplemente el álgebra de todas las matrices $n \times n$. Un álgebra tal que el cuerpo F de los coeficientes sea el de los números complejos y que no contenga ningún elemento propiamente nilpotente, es equivalente a una suma directa de álgebras de matrices totales. Esta muestra de resultados obtenidos por Wedderburn puede dar una idea de las investigaciones realizadas sobre álgebras lineales asociativas generalizadas.

La teoría de anillos e ideales fue edificada sobre una base axiomática y sistematizada por Emmy Noether, una de las pocas grandes mujeres matemáticos que en el mundo han sido, que comenzó a dar clases en Gotinga en 1922. Ya se conocían muchos resultados sobre anillos e ideales cuando ella inició su trabajo, pero consiguió unificarlos en una teoría abstracta a través de una formulación adecuada de los conceptos. Así, por ejemplo, Noether reformula el teorema de la base de Hilbert (cap. 39, sec. 2) de la manera siguiente: un anillo de polinomios en un número cualquiera de indeterminadas sobre un anillo de coeficientes con unidad y una base finita, tiene él mismo una base finita. En esta reformulación, la teoría de invariantes queda integrada dentro del álgebra abstracta.

Emanuel Lasker (1868-1941)[47] había desarrollado ya una teoría de ideales para dominios de polinomios, buscando un método para decidir si un polinomio dado pertenece o no a un ideal generado por otros r polinomios. En 1921[48] demostró Emmy Noether que esta teoría de ideales para polinomios podía deducirse del teorema de la base de Hilbert. De esta manera se establecían unos fundamentos comunes para la teoría de ideales de números enteros algebraicos y de funciones algebraicas enteras (o polinomios). Noether y otros profundizaron mucho más en la teoría abstracta de anillos e ideales, aplicándola a anillos de operadores diferenciales y otras álgebras. Sin embargo, una exposición de los resultados de tales investigaciones nos llevaría demasiado lejos en desarrollos muy especializados.

5. La teoría de álgebras no asociativas

La moderna teoría de anillos o, para ser más exactos, una generalización de la teoría de anillos, incluye también las álgebras no asociativas, en las que la operación producto no es conmutativa ni tampoco asociativa, mientras que las restantes propiedades de las álgebras lineales asociativas siguen siendo válidas. Hoy día se conocen diversos tipos de álgebras no asociativas importantes. El más importante históricamente es el de las álgebras de Lie. En ellas se acostumbra representar el producto de dos elementos a y b por $[a,b]$. En lugar de la propiedad asociativa la operación producto satisface las dos condiciones:

$$[a, b] = -[b, a] \text{ y } [a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0.$$

La segunda propiedad recibe el nombre de identidad de Jacobi. Dicho sea de paso, el producto vectorial de dos vectores cumple las dos condiciones.

Un ideal de un álgebra de Lie L es una subálgebra L_1 tal que el producto de un elemento cualquiera de L por un elemento cualquiera de L_1 pertenece a L_1. Un álgebra de Lie simple es la que no tiene ideales no triviales, y semisimple si no tiene ideales abelianos.

Las álgebras de Lie surgieron de los esfuerzos de Lie por estudiar

[47] Math. Ann., 60, 1905, 20-116.
[48] Math. Ann., 83, 1921, 24-66.

la estructura de sus grupos continuos de transformaciones. Para ello introdujo Lie la idea de transformación infinitesimal.[49] Hablando de una manera intuitiva, una transformación infinitesimal es la que mueve los puntos una distancia infinitesimal. Lie las representó simbólicamente por

$$x_i' = x_i + \delta t X_i(x_1, x_2, ..., x_n), \tag{7}$$

donde δt es una cantidad infinitamente pequeña, o bien

$$\delta x_i = \delta t X_i(x_1, x_2, ..., x_n). \tag{8}$$

La δt es consecuencia de un cambio pequeño en los parámetros del grupo. Así, por ejemplo, supongamos un grupo de transformaciones dado por

$$x_1 = \phi(x, y, a) \quad \text{e} \quad y_1 = \psi(x, y, a)$$

y sea a_0 el valor del parámetro para el que ϕ y ψ definen la transformación identidad, de manera que

$$x = \phi(x, y, a_0) \qquad y = \psi(x, y, a_0).$$

Si incrementamos a_0 a $a_0 + \delta a$, entonces, por el teorema de Taylor

$$x_1 = \phi(x, y, a_0) + \frac{\partial \phi}{\partial a}\delta a + ...$$

$$y_1 = \psi(x, y, a_0) + \frac{\partial \psi}{\partial a}\delta a + ...$$

de manera que despreciando las potencias más altas de δa nos queda

$$\delta x = x_1 - x = \frac{\partial \phi}{\partial a}\delta a \qquad ; \qquad \delta y = y_1 - y = \frac{\partial \psi}{\partial a}\delta a.$$

Para el a_0 fijo, $\partial\phi/\partial a$ y $\partial\psi/\partial a$ son funciones de x e y, de manera que

[49] Archiv fur Mathematik Naturvidenskab, 1, 1876, 152-193 = Ges. Abh., 5, 42-75.

$$\frac{\partial \phi}{\partial a} = \xi(x, y) \quad , \quad \frac{\partial \psi}{\partial a} = \eta(x, y)$$

y

$$\delta x = \xi(x, y)\delta a \quad , \quad \delta y = \eta(x, y)\delta a. \tag{9}$$

Si δa vale δt obtenemos las formas (7) u (8). Las ecuaciones (9) representan una transformación infinitesimal del grupo.

Si $f(x,y)$ es una función analítica de x e y, el efecto sobre ella de una transformación infinitesimal supone reemplazar $f(x,y)$ por $f(x + \xi\, \delta a, y + \eta\, \delta a)$ y aplicando el teorema de Taylor para aproximar hasta el primer orden

$$\delta f = \left(\xi \frac{\partial f}{\partial x} + \eta \frac{\partial f}{\partial y}\right)\delta a.$$

El operador

$$\xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y}$$

es otra manera de *representar* la transformación infinitesimal (9), porque el conocimiento de uno de ellos da el otro. Estos operadores pueden sumarse y multiplicarse en el sentido usual de los operadores diferenciales.

El número de transformaciones infinitesimales independientes, o el número de operadores independientes correspondientes, es el número de parámetros del grupo original de transformaciones. Las transformaciones infinitesimales o los correspondientes operadores, que representaremos por X_1, X_2, ..., X_n, determinan, de hecho, el grupo de Lie de transformaciones. Pero además, y esto es igualmente importante, son ellos mismos generadores de un grupo. Aunque el producto $X_i\, X_j$ no es un operador lineal, la expresión denominada alternante o conmutador de X_i y X_j,

$$X_iX_j - X_jX_i,$$

sí es un operador lineal, y se representa por $[X_i, X_j]$. Con esta operación producto, el grupo de operadores se converte en un álgebra de Lie.

Lie comenzó el trabajo de determinar la estructura de sus grupos (continuos) simples finitos con *r* parámetros, y descubrió cuatro clases principales de álgebras. Wilhelm K. J. Killing (1847-1923) [50] comprobó que esas clases eran correctas para todas las álgebras simples, pero que había además cinco casos excepcionales con 14, 52, 78, 133 y 248 parámetros. El trabajo de Killing no era totalmente riguroso y Elie Cartan intentó cubrir sus huecos.

En su tesis *Sur la structure des groupes de transformations finis et continus*,[51] da Cartan una clasificación completa de todas las álgebras de Lie simples sobre el cuerpo de los complejos para los parámetros y las variables. Tal como había obtenido Killing, Cartan descubrió que se dividían en cuatro casos generales y las cinco álgebras excepcionales. Cartan construyó explícitamente estas cinco álgebras excepcionales. En 1914 [52] determinó Cartan todas las álgebras simples con valores reales para los parámetros y las variables. Sus resultados son aún básicos.

El uso de representaciones para estudiar álgebras de Lie ha resultado muy eficaz, tal como en el caso de los grupos abstractos. En su tesis, y en un artículo de 1913,[53] encontró Cartan representaciones irreducibles de las álgebras de Lie simples. Más tarde obtendría Hermann Weyl un resultado clave[54]: cualquier representación de un álgebra de Lie semisimple (sobre un cuerpo algebraicamente cerrado de característica cero) es completamente reducible.

6. Panorama general del álgebra abstracta

Nuestras breves indicaciones de los resultados obtenidos en el dominio del álgebra abstracta, ciertamente no dan una imagen completa de los desarrollos ni siquiera del primer cuarto de nuestro siglo. Puede ser útil, sin embargo, mostrar el vasto campo que se ha abierto por el uso sistemático de un enfoque abstracto.

---

[50] Math. Ann., 31, 1888, 252-290 y vols. 33, 34 y 36.

[51] 1894; Paris, Vuiber, 2.ª ed., 1933 = *Œuvres*, parte I, vol. 1, 137-286.

[52] Ann. de l'Ecole Norm. Sup., 31, 1914, 263-355 = *Œuvres*, parte I, vol. 1, 399-491.

[53] Bull. Soc. Math. de France, 41, 1913, 53-96 = *Œuvres*, parte I, vol. 1, 355-398.

[54] Mathematische Zeitschrift, 23, 1925, 271-309 y 24, 1926, 328-395 = *Ges. Abh.*, 2, 543-647.

Hasta 1900 más o menos, las diversas cuestiones de tipo algebraico estudiadas, ya fueran matrices, álgebras de formas en dos, tres o *n* variables, hipernúmeros, congruencias o la resolución de ecuaciones polinómicas, habían estado basadas en los sistemas de números reales o complejos. El movimiento algebraico abstracto introdujo, sin embargo, los grupos, anillos, ideales, álgebras con división y cuerpos abstractos. Aparte de investigar las propiedades de tales estructuras abstractas y relaciones como las de isomorfismo y homomorfismo, los matemáticos han descubierto la posibilidad de tomar casi cualquier cuestión algebraica y plantearse problemas sobre ella reemplazando los números reales o complejos por cualquier otra estructura abstracta. Por ejemplo, en lugar de matrices con elementos complejos, pueden estudiarse matrices con elementos de un anillo o un cuerpo arbitrario. Análogamente se pueden considerar problemas de la teoría de números y, sustituyendo los números enteros por un anillo cualquiera, reconsiderar cualquier cuestión que antes se había investigado solamente para los enteros usuales. Pueden considerarse incluso funciones y series de potencias con coeficientes en un cuerpo arbitrario.

Todas estas generalizaciones se han hecho realmente. Ya hemos visto que Wedderburn, en su trabajo de 1907, generalizó resultados anteriores sobre álgebras lineales asociativas (hipernúmeros) reemplazando los coeficientes reales o complejos por otros de un cuerpo cualquiera. Pues bien, es posible sustituir el cuerpo por un anillo e investigar los teoremas que siguen verificándose pese a este cambio. Se ha estudiado incluso la teoría de ecuaciones con coeficientes en un cuerpo arbitrario, incluso finito.

Como otro ejemplo de la moderna tendencia a generalizar, piénsese en las formas cuadráticas. Las formas cuadráticas con coeficientes enteros fueron importantes en el estudio de la representación de enteros como sumas de cuadrados, y con coeficientes reales en el estudio de las cónicas y las cuádricas. En el siglo XX se han estudiado las formas cuadráticas con coeficientes en cualquier cuerpo. Según se van introduciendo más estructuras abstractas, pueden ser utilizadas como base o dominio de coeficientes de otras teorías algebraicas más antiguas, y el proceso de generalización avanza indefinidamente. Este uso de conceptos abstractos exige a su vez el uso de técnicas algebraicas también abstractas; así, muchos campos anteriormente no relacionados entre sí se han visto absorbidos por el álgebra abstracta. Este ha sido el caso, en particular, con extensas partes de la

teoría de números, incluida también la teoría de números algebraicos.

Sin embargo, el álgebra abstracta ha terminado por subvertir su propio papel dentro de la matemática. Sus conceptos se formularon para unificar dominios matemáticos aparentemente diversos y completamente separados, tal como hizo, por ejemplo, la teoría de grupos. Una vez formuladas las teorías abstractas, los matemáticos olvidaron los campos concretos originales y concentraron su atención únicamente en las estructuras abstractas. Con la introducción de cientos de conceptos subordinados, la materia se ha desarrollado como los hongos en un desorden de desarrollos menores que tienen poca relación unos con otros y con los campos concretos originales. La unificación ha cedido su lugar a la diversificación y a la especialización. En realidad, la mayor parte de los investigadores actuales en álgebra abstracta ignoran incluso los orígenes de las estructuras abstractas, y tampoco están interesados en las aplicaciones de sus resultados a campos concretos.

Bibliografía

Artin, Emil: «The Influence of J. H. M. Wedderburn on the Development of Modern Algebra». *Amer. Math. Soc. Bull.*, 56, 1950, 65-72.

Bell, Eric T.: «Fifty Years of Algebra in America, 1888-1938». *Amer. Math. Soc. Semicentennial Publications*, II, 1938, 1-34.

Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.

Cartan, Elie: «Notice sur les travaux scientifiques». *Œuvres complètes*, Gauthier-Villars, 1952-1955, part I, vol. I, 1-98.

Dicke, Auguste: *Emmy Noether, 1882-1935*, Birkhäuser Verlag, 1970.

Dickson, L. E.: «An Elementary Exposition of Frobenius's Theory of Group-Characters and Group-Determinants». *Annals of Math.*, 4, 1902, 25-49.

— *Linear Algebras*, Cambridge University Press, 1914.

— *Algebras and Their Arithmetics* (1923), G. E. Stechert (reprint), 1938.

Frobenius, F. G.: *Gesammelte Abhandlungen*, 3 vols., Springer-Verlag, 1968.

Hawkins, Thomas: «The Origins of the Theory of Group Characters». *Archive for History of Exact Sciences*, 7, 1971, 142-170.

MacLane, Saunders: «Some Recent Advances in Algebra». *Amer. Math. Monthly*, 46, 1939, 3-19. Also in Albert, A. A., ed.: *Studies in Modern Algebra*, The Math. Assn. of Amer., 1963, 9-34.

— «Some Additional Advances in Algebra» in Albert, A. A., ed.: *Studies in Modern Algebra,* The Math. Assn. of Amer., 1963, 35-58.
Ore, Oystein: «Some Recent Developments in Abstract Algebra». *Amer. Math. Soc. Bull.,* 37, 1931, 537-548.
— «Abstract Ideal Theory». *Amer. Math. Soc. Bull.,* 39, 1933, 728-745.
Steinitz, Ernst: *Algebraische Theorie der Körper,* W. de Gruyter, 1910; 2.ª ed., 1930; Chelsea (reprint), 1950. La primera edición es la misma que el artículo en *Jour. für Math.,* 137, 1910, 167-309.
Wiman, A.: «Endliche Gruppen linearer Substitutionen». *Encyk. der Math. Wiss.,* B. G. Teubner, 1898-1904, I, part 1, 522-554.
Wussing, H. L.: *Die Genesis des abstrakten Gruppenbegriffes,* VEB Deutscher Verlag der Wissenschaften, 1969.

Capítulo 50

LOS ORIGENES DE LA TOPOLOGIA

> Creo que nos falta otro tipo de análisis propiamente geométrico o lineal que exprese directamente la localización, así como el álgebra expresa la magnitud.
>
> G. W. LEIBNIZ

1. ¿Qué es la topología?

Cierto número de desarrollos durante el siglo XIX cristalizaron en una nueva rama de la geometría, que hoy llamamos topología pero conocida durante largo tiempo con el nombre de «análisis situs». Para decirlo de una manera vaga, por el momento, la topología se ocupa de aquellas propiedades de las figuras que permanecen invariantes cuando dichas figuras son plegadas, dilatadas, contraídas o deformadas de cualquier manera tal que no aparezcan nuevos puntos o se hagan coincidir puntos ya existentes. La transformación permitida presupone, en otras palabras, que hay una correspondencia biunívoca entre los puntos de la figura original y los de la figura transformada, y que la transformación hace corresponder puntos próximos a puntos próximos. Esta última propiedad se llama continuidad, y lo que se requiere es que la transformación y su inversa sean ambas continuas. Una transformación así se llama un homeomorfismo. La topología se describe a veces de una manera imprecisa como la geometría de las bandas de goma, porque si las figuras fueran de goma sería posible deformar muchas de ellas en otras figuras homeomorfas. Así una banda de goma plana puede deformarse para convertirse

en, y es por tanto topológicamente lo mismo que, un círculo o un cuadrado, pero no es topológicamente lo mismo que una figura en forma de ocho, porque esto requeriría la superposición de dos puntos de la banda.

Se suele imaginar las figuras como sumergidas en un espacio ambiente. A efectos topológicos dos figuras pueden ser homeomorfas incluso aunque no sea posible transformar topológicamente todo el espacio en el que está sumergida la primera figura en el espacio que contiene a la segunda. Por ejemplo, si se toma un rectángulo de papel alargado y se unen los dos lados cortos, se obtiene una superficie cilíndrica. Si en lugar de esto se gira uno de los lados cortos 360º y después se unen los dos lados cortos, la nueva figura obtenida es topológicamente equivalente a la anterior; sin embargo, es imposible transformar topológicamente el espacio tridimensional en sí mismo de tal manera que la primera figura se transforme en la segunda.

La topología, tal como la concebimos hoy, se divide en dos ramas separadas, en cierto sentido: la topología conjuntista (o topología general), que se ocupa de las figuras geométricas consideradas como conjuntos de puntos, donde el conjunto total es frecuentemente considerado como un espacio; y la topología combinatoria o algebraica, que trata las figuras como agregados de bloques más pequeños, exactamente lo mismo que una pared es una colección de ladrillos. Desde luego, los conceptos de la topología de conjuntos de puntos se usan también en la topología combinatoria, especialmente para estructuras geométricas muy generales.

La topología ha tenido numerosos y variados orígenes. Como en la mayor parte de las ramas de la matemática, se recorrieron muchas etapas que sólo posteriormente fueron consideradas como perteneciendo a (o capaces de ser englobadas en) una nueva teoría. En el caso que nos ocupa, la posibilidad de un estudio significativo por separado, fue al menos esbozada por Klein en su *Erlanger Programm* (cap. 38, sec. 5). Klein estaba generalizando los tipos de transformaciones estudiadas en geometría proyectiva y geometría algebraica, y ya era consciente de la importancia de los homeomorfismos por la obra de Riemann.

2. La topología conjuntista

La teoría de conjuntos de puntos tal como fue iniciada por Cantor (cap. 41, sec. 7), y continuada por Jordan, Borel y Lebesgue (cap. 44, secs. 3 y 4) no se refiere *eo ipso* a transformaciones y propiedades topológicas. Pero por otra parte, cualquier conjunto de puntos considerado como un espacio es ya de interés para la topología. Lo que distingue a un espacio de un simple conjunto de puntos es algún concepto que ligue unos puntos con otros. Así, en un espacio euclídeo, la distancia entre puntos nos dice cómo están de próximos unos a otros, y en particular nos permite definir qué es un punto límite de un conjunto.

Los orígenes de la topología conjuntista ya han sido mencionados (cap. 46, sec. 2). Fréchet, el año 1906, impulsado por el deseo de unificar la teoría de conjuntos de puntos de Cantor y el tratamiento de las funciones como puntos de un espacio, que ya se había hecho habitual en el cálculo de variaciones, inició el estudio de los espacios abstractos. La aparición del análisis funcional con la introducción de los espacios de Hilbert y de Banach dio una importancia adicional al estudio de los conjuntos de puntos como espacios. Las propiedades que se mostraron relevantes para el análisis funcional son topológicas, en gran medida por la importancia de los límites de sucesiones. Por otra parte, los operadores del análisis funcional son simplemente transformaciones de un espacio en otro.[1]

Como hizo notar Fréchet, la propiedad que liga los puntos unos con otros no necesita ser la distancia euclídea. El introdujo (cap. 46, sec. 2) varios conceptos diferentes que pueden utilizarse para indicar cuándo un punto es límite de una sucesión de puntos. En particular, generalizó la idea de distancia introduciendo la clase de los espacios métricos. En un espacio métrico, por ejemplo, el espacio euclídeo bidimensional, uno habla de un entorno de un punto entendiendo por ello el conjunto de todos los puntos cuya distancia al dado es menor que un cierto número ε. Tales entornos son los entornos circulares; de manera análoga se podrían usar entornos cuadrados. Sin embargo, es posible también suponer que los entornos, ciertos

[1] Las definiciones de algunas propiedades básicas de los conjuntos de puntos, tales como la compacidad y la separabilidad, han sido distintas según los diferentes autores y todavía no han sido unificadas. Aquí utilizaremos los significados usados más comúnmente en la actualidad.

subconjuntos de un conjunto de puntos dado, vengan definidos de alguna otra manera, incluso sin la introducción previa de una métrica. Se dice entonces que tales espacios tienen una topología de entornos; este concepto es una generalización de los espacios métricos. Felix Hausdorff (1868-1942), en su «*Grundzüge der Mengenlehre*» (Fundamentos de Teoría de Conjuntos, 1914) utilizó el concepto de entorno (que ya había usado Hilbert en 1902 en una formulación axiomática especial de la geometría euclídea plana) y edificó una teoría definitiva de los espacios abstractos basada en este concepto.

Hausdorff define un espacio topológico como un conjunto de elementos x, junto con una familia de subconjuntos U_x asociada a cada punto x. Estos subconjuntos se llaman entornos y deben satisfacer las condiciones siguientes:

a) Para cada punto x existe el menos un entorno U_x que lo contiene.

b) La intersección de dos entornos cualesquiera de x contiene otro entorno de x.

c) Si y es un punto que pertenece a U_x, entonces existe un U_y tal que $U_y \subseteq U_x$.

a) Si $x \neq y$, entonces existen U_y y U_x tales que $U_x \cap U_y = \emptyset$.

Hausdorff introdujo también los axiomas de numerabilidad:

a) Para todo punto x el conjunto de los U_x es como máximo numerable.

b) El conjunto de todos los entornos distintos es numerable.

La tarea básica en topología conjuntista consiste en definir diversas nociones fundamentales. Por ejemplo, un punto límite de un conjunto de puntos en un espacio de entornos es un punto tal que todo entorno suyo contiene otros puntos del conjunto; un conjunto es abierto si todo punto perteneciente a él puede ser encerrado en un entorno que contiene sólo puntos del conjunto; si un conjunto contiene a todos sus puntos límites entonces se llama cerrado; un espacio o un subconjunto se llama compacto si todo subconjunto infinito suyo tiene un punto límite. Así, por ejemplo, la recta euclídea usual no es un conjunto compacto porque el subconjunto infinito formado por los puntos de abscisa un número entero, no tiene ningún punto límite. Un conjunto es conexo si, de cualquier manera que se le divida en dos subconjuntos disjuntos, uno de ellos al menos contiene puntos límites del otro; la curva $y = \operatorname{tg} x$ no es conexa, pero la curva $y = \operatorname{sen} 1/x$ más el intervalo $(-1,1)$ del eje Y es conexa.

La separabilidad, introducida por Fréchet en su tesis de 1906, es otro concepto básico; un espacio se denomina separable si tiene un subconjunto numerable cuya adherencia (el conjunto más sus puntos límites) coincide con el espacio total.

Ahora ya podemos introducir los conceptos de transformación continua y homeomorfismo. Una transformación continua es una correspondencia que a cada punto de un espacio asocia un único punto de un segundo espacio o espacio imagen, y tal que dado un entorno arbitrario de un punto imagen existe un entorno del punto original (o de cada punto original si hay más de uno) cuyos puntos se transforman en puntos del entorno dado en el espacio imagen. Este concepto no es más que una generalización de la definición ε-δ de una función continua, donde el ε determina el entorno del punto en el espacio imagen y el δ un entorno del punto original. Un homeomorfismo entre dos espacios S y T es una correspondencia biunívoca que es continua en los dos sentidos, es decir, la transformación de S a T y la de T a S son continuas. La tarea básica de la topología general es entonces la de descubrir propiedades que sean invariantes bajo transformaciones continuas y homeomorfismos. Todas las propiedades mencionadas más arriba son invariantes topológicos.

Hausdorff contribuyó con muchos resultados a la teoría de espacios métricos. En particular desarrolló el concepto de completitud que había introducido Fréchet en su tesis de 1906. Un espacio se llama completo si toda sucesión $\{a_n\}$ que cumpla la condición de que dado un ε positivo exista un N tal que $|a_n - a_m| < \varepsilon$ para todos m y n mayores que N, tiene límite. Hausdorff demostró que todo espacio métrico puede extenderse a un espacio métrico completo de una y sólo una manera.

La introducción de los espacios abstractos planteó diversos problemas que provocaron una intensa investigación. Por ejemplo, si un espacio está definido por medio de entornos ¿es necesariamente metrizable?; es decir, ¿es posible introducir una métrica que preserve la estructura del espacio de tal manera que los puntos límites permanezcan como puntos límites? Este problema fue planteado por Fréchet. Un resultado debido a Paul S. Urysohn (1898-1924), afirma que todo espacio normal es metrizable;[2] un espacio normal es aquel en que dos conjuntos cerrados disjuntos cualesquiera pueden ser

[2] Math. Ann., 94, 1925, 262-295.

separados por dos abiertos disjuntos. Otro resultado importante relacionado con éste y debido también a Urysohn es el siguiente:[3] todo espacio métrico separable, es decir, todo espacio métrico que contenga un subconjunto denso numerable en el espacio, es homeomorfo a un subconjunto del cubo de Hilbert; el cubo de Hilbert consiste en el espacio de todas las sucesiones $\{x_i\}$ de números reales tales que $0 \leqslant x_i \leqslant 1/i$ y donde la distancia se define como

$$d = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2}.$$

El problema de la dimensión, ya mencionado, apareció por la demostración de Cantor de la existencia de una correspondencia biunívoca entre la recta y el plano (cap. 41, sec. 7) y por la construcción de la curva de Peano que llena un cuadrado (cap. 42, sec. 5). Fréchet (que trabajaba ya con espacios abstractos) y Poincaré vieron la necesidad de una definición de dimensión que se aplicase también a los espacios abstractos y que, sin embargo, atribuyese a la recta y al plano las dimensiones usuales. La definición que había sido aceptada tácitamente hasta entonces era la del número de coordenadas necesarias para determinar los puntos del espacio; esta definición no era aplicable a los espacios generales.

En 1912 [4] Poincaré dio una definición recursiva. Un continuo (o conjunto cerrado conexo) tiene dimensión n si puede ser separado en dos partes cuya frontera común consiste en continuos de dimensión $n-1$. Luitzen E. J. Brouwer (1881-1967) hizo notar que esta definición no era aplicable al cono de dos hojas, porque las hojas se separan por un punto. La definición de Poincaré fue mejorada por Brouwer,[5] Urysohn [6] y Karl Menger (n. 1902).[7]

Las definiciones de Menger y Urysohn son análogas y a ambos se les atribuye la definición generalmente aceptada hoy. Su definición asigna una dimensión local. La formulación de Menger es la siguiente: el conjunto vacío tendría por definición dimensión -1. Un conjunto M se llama n-dimensional en un punto P si n es el

[3] Math. Ann., 94, 1925, 309-315.

[4] Revue de Métaphysique et de Morale, 20, 1912, 483-504.

[5] Jour. für Math., 142, 1913, 146-152.

[6] Fundamenta Mathematicae, 7, 1925, 30-137 y 8, 1926, 225-359.

[7] Monatshefte für Mathematik und Physik, 33, 1923, 148-160 y 34, 1926, 137-161.

mínimo número para el cual existen entornos de P arbitrariamente pequeños cuyas fronteras en M tengan dimensión menor que n. El conjunto M se llama n-dimensional si su dimensión es menor o igual que n en cada uno de sus puntos, pero igual a n en un punto al menos.

Otra definición ampliamente aceptada es la de Lebesgue;[8] un espacio es n-dimensional si n es el número mínimo para el cual los recubrimientos cualesquiera mediante conjuntos cerrados de diámetro arbitrariamente pequeño contienen algún punto común a $n + 1$ conjuntos del recubrimiento. Los espacios euclídeos tienen la dimensión correcta con cualquiera de estas definiciones, y la dimensión de un espacio arbitrario es un invariante topológico.

Un resultado clave en teoría de la dimensión es el teorema debido a Menger (*«Dimensionstheorie»*, 1928, p. 295) y a A. Georg Nöbeling (1907),[9] que afirma que todo espacio métrico compacto n-dimensional es homeomorfo a un subconjunto del espacio euclídeo $(2n + 1)$-dimensional.

Otro problema planteado por los trabajos de Jordan y Peano era la definición misma de curva (cap. 42, sec. 5). La respuesta fue posible basándose en resultados de teoría de la dimensión. Menger [10] y Urysohn [11] definieron una curva como un continuo unidimensional, entendiendo por continuo un conjunto de puntos cerrado y conexo. (Esta definición requiere que una curva abierta, como una parábola, se cierre mediante un punto del infinito.) Esta definición excluye las curvas que llenan un espacio y hace de la propiedad de ser una curva un invariante bajo homeomorfismos.

El campo de la topología general o conjuntista ha seguido siendo extraordinariamente activo. Es relativamente fácil aquí el introducir variaciones, particularizar y generalizar las caracterizaciones axiomáticas de los diversos tipos de espacios. Se han introducido cientos de definiciones y de teoremas, aunque en la mayor parte de los casos el valor último de estos conceptos es dudoso. Como en otros campos, los matemáticos no han dudado en zambullirse libremente y sin complejos en la topología general.

---

[8] Fundamenta Mathematicae, 2, 1921, 256-285.

[9] Math. Ann., 104, 1930, 71-80.

[10] Monatshefte für Mathematik und Physik, 33, 1923, 148-160 y Math. Ann., 95, 1926, 266-306.

[11] Fundamenta Mathematicae, 7, 1925, 30-137, parte de la p. 93.

3. Los comienzos de la topología combinatoria

Tan pronto como en 1679 intentó ya Leibniz, en su «*Characteristica Geometrica*», formular las propiedades básicas de las figuras geométricas, utilizar símbolos especiales para representarlas sintéticamente y combinar estas propiedades mediante ciertas operaciones para producir otras. Leibniz llamó a este estudio análisis situs o *geometria situs*. En una carta a Huygens de 1679 [12] explica que no le satisface el tratamiento de las figuras geométricas mediante coordenadas porque, aparte del hecho de que este método no era directo ni bello, se refería a magnitudes, mientras que «Creo que nos falta otro tipo de análisis propiamente geométrico o lineal que exprese directamente la localización *(situs)*, así como el álgebra expresa la magnitud». Los pocos ejemplos que da Leibniz acerca de lo que se proponía construir implicaban aún propiedades métricas, aunque él aspiraba a conseguir unos algoritmos geométricos que suministrasen la solución de los problemas puramente geométricos. Debido probablemente a que Leibniz era más bien impreciso acerca del tipo de geometría que buscaba, Huygens no se mostró muy entusiasmado con sus ideas y su simbolismo. En la medida en que las cosas estaban medianamente claras, lo que Leibniz preveía era lo que llamamos hoy topología combinatoria.

Una propiedad combinatoria de ciertas figuras geométricas es la siguiente, atribuida a Euler aunque era conocida ya por Descartes en 1639 y, a través de los manuscritos no publicados de este último, por Leibniz en 1675: si se cuentan los números de vértices, caras y aristas de un poliedro convexo cualquiera, por ejemplo un cubo, entonces se tiene que $V - A + C = 2$. Este hecho fue publicado por Euler en 1750 [13] y en 1751 presentó una demostración.[14] Euler estaba interesado en esta relación para utilizarla al clasificar los poliedros; aunque había descubierto una propiedad de todos los poliedros cerrados convexos, no pensó en su posible invariancia bajo

---

[12] Leibniz, *Math. Schriften*, 1 Abt., vol. 2, 1850, 19-20 = Gerhardt, *Der Briefwechsel von Leibniz mit Mathematikern*, 1, 1899, 568 = Chr. Huygens, *Oeuv. Comp.*, 8, n.º 2192.

[13] Novi Comm. Acad. Sci. Petrop., 4, 1752-1753, 109-140, pub. 1758 = *Opera*, (1), 26, 71-93.

[14] Novi Comm. Acad. Sci. Petrop., 4, 1752-1753, 140-160, pub. 1758 = *Opera*, (1), 26, 94-108.

transformaciones continuas, ni tampoco definió la clase de los poliedros para los que dicha relación se verifica.

En 1811 Cauchy [15] dio otra demostración suprimiendo el interior de una cara y desarrollando la figura restante sobre un plano; de esta forma se obtiene un polígono para el cual $V - A + C$ debe valer 1. Cauchy demostraba esto último triangulando la figura y contando después los cambios cuando los triángulos se van suprimiendo uno por uno. Esta demostración, a pesar de ser incorrecta porque supone que cualquier poliedro convexo cerrado es homeomorfo a una esfera, fue aceptada sin reparos por los matemáticos del siglo XIX.

Otro problema muy conocido, que fue en su tiempo una curiosidad, pero cuya naturaleza topológica fue reconocida más tarde, es el problema de los puentes de Könisberg.

Sobre el Pregel, río que atraviesa Könisberg, hay construidos siete puentes que unen las dos islas con las riberas, puentes señalados con la letra *b* en la figura 50.1. Los habitantes de la ciudad se divertían intentando cruzar los siete puentes en un paseo continuo sin cruzar dos veces ninguno de ellos. Euler, que vivía entonces en San Petersburgo, oyó hablar del problema y lo resolvió en 1735 [16] simplificando su planteamiento al reemplazar las zonas de tierra por puntos y los puentes por segmentos o arcos, tal como aparece en la figura 50.2; la cuestión que Euler se planteó entonces fue la de si

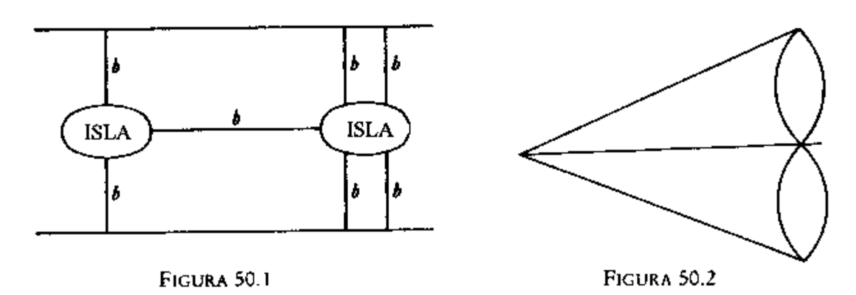


FIGURA 50.1 FIGURA 50.2

[15] Jour. de L'Ecole Poly., 9, 1813, 68-86 y 87-98 = Œuvres, (2), 1, 7-38.

[16] Comm. Acad. Sci. Petrop., 8, 1736, 128-140, pub. 1741. Puede encontrarse una traducción inglesa de este artículo en James R. Newman, *The world of mathematics*, Simon and Schuster, 1956, vol. 1, 573-580. Hay versión castellana, *Sigma. El mundo de las matemáticas*, Barcelona, Grijalbo, 8.ª ed., 1983.

era posible describir esta figura en un movimiento continuo del lápiz sin recorrer dos veces ninguno de los trazos. Demostró que no era posible en este caso, y dio un criterio acerca de cuándo tales caminos son o no posibles para conjuntos dados de puntos y arcos.

Gauss insistió a menudo [17] en la necesidad de estudiar las propiedades geométricas básicas de las figuras, pero no hizo ninguna contribución sobresaliente sobre el tema. En 1848, Johann B. Listing (1806-1882), alumno de Gauss en 1834 y más tarde profesor de física en Göttingen, publicó su *Vorstudien zur Topologie*, donde discutía lo que él hubiera preferido llamar geometría de posición, pero que denominó topología, ya que el primer nombre había sido utilizado el año anterior por Von Staudt para denominar la geometría proyectiva. En 1858 comenzó una nueva serie de investigaciones topológicas que aparecieron publicadas bajo el título de *Der Census raümlicher Complexe* (Panorama de los Complejos Espaciales).[18] Listing buscaba leyes cualitativas para las figuras geométricas; por ejemplo, trató de generalizar la relación $V - A + C = 2$ de Euler.

El primero que formuló correctamente el carácter de las investigaciones topológicas fue Möbius, que había sido ayudante de Gauss en 1813. Möbius había clasificado ya diversos tipos de propiedades geométricas, proyectivas, afines, semejanzas y congruencias, y en 1863, en su *Theorie der elementaren Verwandschaft* (Teoría de las relaciones elementales) [19] se propuso estudiar la relación existente entre dos figuras cuyos puntos están en una correspondencia biunívoca tal que puntos próximos corresponden a puntos próximos. Comenzó estudiando la *geometría situs* de los poliedros, subrayando que un poliedro puede ser considerado como una colección de polígonos bidimensionales, los cuales, triangulados, convertirían al poliedro en una colección de triángulos. Esta idea resultó ser fundamental. También demostró [20] que algunas superficies podían cortarse y desarrollarse como polígonos, identificando adecuadamente ciertos lados. Así, un doble anillo podía ser representado como un polígono (fig. 50.3) identificando los lados que están señalados con las mismas letras.

---

[17] *Werke*, 8, 270-286.

[18] Abh. der Ges. der Wiss. zu Gött., 10, 1861, 97-180, y como libro en 1862.

[19] Königlich Sachsischen Ges. der Wiss. zu Leipzig, 15, 1863, 18-57 = *Werke*, 2, 433-471.

[20] *Werke*, 2, 518-559.

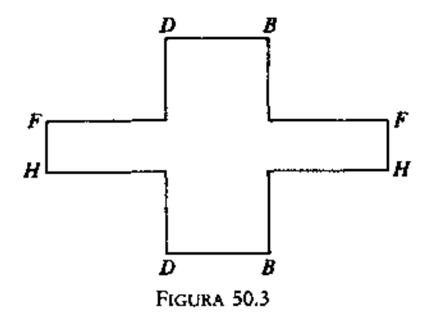


FIGURA 50.3

En 1858 Möbius y Listing descubrieron independientemente las superficies de una sola cara, de las cuales la banda de Möbius es la más conocida (fig. 50.4); esta figura se obtiene cogiendo una banda rectangular de papel, girando uno de sus extremos cortos 180° y uniéndolo al extremo opuesto.

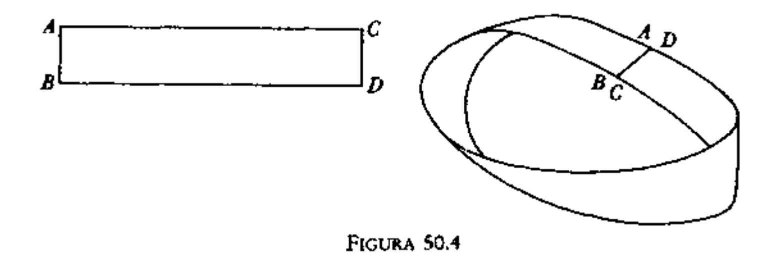


FIGURA 50.4

Listing publicó su descubrimiento en *Der Census*, y la banda en cuestión también aparece descrita en una publicación de Möbius.[21] Por lo que se refiere a esta banda, el tener una sola cara puede ser caracterizado por el hecho de que se la puede pintar totalmente con un movimiento continuo de la brocha. Si se pretende, en cambio, pintar una banda que no ha sido previamente girada, la brocha debe saltar sobre uno de sus bordes para pasar de una cara a la otra. El hecho de tener una sola cara puede caracterizarse también utilizando un vector perpendicular a la superficie; si al desplazar este vector

---

[21] Königlich Sachsischen der Wiss. zu Leipzig, 17, 1865, 31-68 = *Werke*, 2, 473-512; ver también p. 519.

arbitrariamente sobre ella, manteniéndolo siempre perpendicular a la superficie, conserva el mismo sentido cuando vuelve a su posición original, entonces se dice que dicha superficie tiene dos caras, y si el sentido se invierte, que tiene una sola cara. Sobre la banda de Möbius el vector perpendicular regresa al punto de partida pero sobre el «lado opuesto», es decir, con su sentido invertido.

Otro problema que más tarde mostró su carácter topológico es el llamado problema del mapa. Consiste en demostrar que cuatro colores son suficientes para colorear cualquier mapa, de manera que países que tengan por lo menos un arco de frontera común estén coloreados con colores distintos. Francis Guthrie (m. 1899), profesor de matemáticas poco conocido, formuló por primera vez en 1852 la conjetura de que cuatro colores siempre serían suficientes, y por aquella época su hermano Frederick se lo comunicó a De Morgan. El primer artículo dedicado a este problema fue el de Cayley de 1879;[22] en él decía que no había podido conseguir una demostración de dicha conjetura. Muchos matemáticos han seguido buscando desde entonces una demostración, y aunque algunas de ellas fueron aceptadas durante cierto tiempo, se vio posteriormente que todas eran incorrectas y el problema aún sigue abierto (*).

El mayor impulso a las investigaciones topológicas vino de los trabajos de Riemann en teoría de funciones de variable compleja. En su tesis de 1851 sobre funciones de variable compleja y en sus estudios sobre funciones abelianas,[23] insiste en que para trabajar con funciones eran indispensables algunos teoremas del «análisis situs». En estas investigaciones se encontró con la necesidad de introducir la idea de conexión sobre las superficies de Riemann, definiéndola de la manera siguiente: «Si sobre la superficie F (con frontera) pueden dibujarse n curvas cerradas $a_1, a_2, ..., a_n$ que ni individualmente, ni combinadas, limitan completamente una parte de dicha superficie F, pero tales que con ayuda de ellas cualquier otra curva cerrada forma la frontera completa de una parte de F, entonces la superficie

---

[22] Proceedings of the Royal Geographic Society, 1, 1879, 259-261 = *Coll. Math. Papers*, 11, 7-8.

[23] Jour. für Math., 54,1857, 105-110 = *Werke*, 91-96, ver también *Werke*,479-482.

(*) El problema de los cuatro colores ha sido resuelto por Appel y Haken utilizando computadores. Véase K. Appel y W. Haken, Every planar map is four colorable, I: Discharching, Illinois J. Math., 21 (1977), 429-490, y K. Appel, W. Haken y J. Koch, Every planar map is four colorable, II: Reducibility, *ibid.*, 491-567. (Nota del traductor.)

se llama ($n + 1$)-conexa». Para reducir la conexión de una superficie (con frontera) establece Riemann:

Por medio de una sección *[Querschnitt]*, es decir, de una línea sobre la superficie que vaya de un punto frontera a un punto frontera, una superficie ($n + 1$)-conexa puede transformarse en otra n-conexa F'. Las partes de la frontera que aparecen en esta división hacen el papel de frontera incluso durante las divisiones posteriores, de manera que una sección no puede pasar más de una vez por un mismo punto pero puede terminar en uno de sus puntos previos.

... Para aplicar estas consideraciones a una superficie sin frontera, cerrada, tenemos primero que transformarla en una con frontera destacando un punto arbitrario y formando la primera división mediante este punto y una sección que comienza y termina en él, es decir, por una curva cerrada.

Riemann pone el ejemplo de un toro (fig. 50.5) que es una superficie 3-conexa (género 1 o número de Betti unidimensional 2) y que puede transformarse en una superficie simplemente conexa por medio de una curva cerrada *abc* y una sección *ab'c'*.

Riemann había clasificado de esta forma las superficies según su tipo de conexión y, como él mismo constató, había introducido así una propiedad topológica. En términos del género, nombre usado por los geómetras algebraicos de finales del siglo XIX, Riemann había clasificado las superficies cerradas por medio de su género p, siendo $2p$ el número de curvas cerradas (secciones-lazo o *«Rückkehrschnitte»)* necesarias para hacer la superficie simplemente conexa, y $2p + 1$ para cortar la superficie en dos partes distintas. Riemann consideraba como intuitivamente evidente que si dos superficies de Riemann cerradas (orientables) son topológicamente equivalentes, entonces tienen el mismo género. Observó también que todas las superficies algebraicas cerradas de género cero, es decir, simplemente conexas, son topológicamente (y conformemente y birracionalmente) equivalentes, pudiéndose aplicar topológicamente sobre una esfera.

Debido a que la estructura de las superficies de Riemann es complicada y a que figuras topológicamente equivalentes tienen el mismo género, algunos matemáticos se dedicaron a buscar estructuras más sencillas. William K. Clifford demostró [24] que la superficie de Riemann de una función n-valuada con w puntos de ramificación, puede

[24] Proc. London Math. soc., 8, 1877, 292-304 = *Math. Papers*, 241-254.

transformarse en una esfera con p agujeros, siendo $p = (w/2) - n + 1$ (fig. 50.6). Es probable que Riemann ya conociera y usara este modelo. Klein sugirió otro modelo topológico, la esfera con p asas (fig. 50.7).[25]

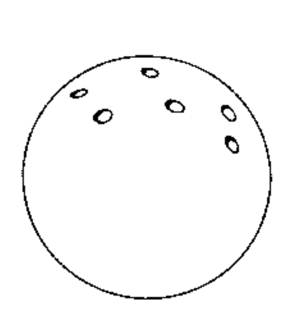


FIGURA 50.5

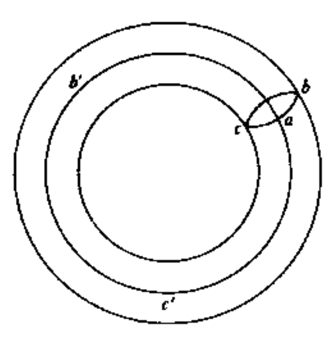


FIGURA 50.6

El estudio de la equivalencia topológica de las superficies cerradas se debe a un gran número de matemáticos. Para exponer el resultado principal es necesario dar la definición de superficie orientable: una superficie orientable es aquella que se puede triangular de manera que cada triángulo (curvilíneo) puede ser orientado de tal forma que cada lado común a dos triángulos tiene orientaciones opuestas inducidas sobre él por las de los dos triángulos. Así, por ejemplo, la esfera es orientable pero el plano proyectivo (ver más adelante) no lo es, lo cual fue descubierto por Klein.[26] El resultado fundamental, clarificado por Klein en su trabajo, es que dos superficies cerradas orientables son homeomorfas si y sólo si tienen el mismo género. Para las superficies orientables con frontera, debe añadirse a la condición anterior el que el número de curvas frontera sea el mismo en ambas; este teorema había sido demostrado ya por Jordan.[27]

---

[25] *Uber Riemanns Theorie der algebraischen Funktionen und ihrer Integrale*, B. G. Teubner, 1882; reimpresión inglesa en Dover, 1963. También en Klein, *Ges. Math. Abh.*, 3, 499-573.

[26] Math. Ann., 7, 1874, 549-557 = *Ges. Math. Abh.*, 2, 63-77.

[27] Jour. für Math., (2), 11, 1866, 105-109 = *Œuvres*, 4, 85-89.

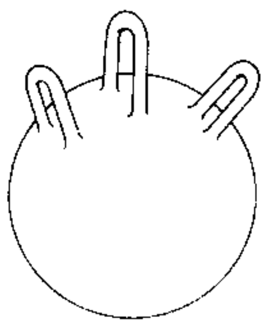


FIGURA 50.7

La complejidad que pueden presentar las figuras cerradas, incluso las bidimensionales, fue puesta de relieve por Klein al introducir en 1882 la superficie llamada posteriormente botella de Klein (fig. 50.8) (véase la sec. 23 de la referencia en la nota 25). El cuello de esta botella penetra en su interior sin atravesarla y termina unido de una manera continua a la base a lo largo de *C*. A lo largo de *D* la superficie no se interrumpe y sin embargo el cuello penetra en el interior de la superficie; la superficie no tiene aristas, ni interior ni exterior; tiene una sola cara y número de conexión unidimensional 3, o bien género igual a 1; no puede ser construida en tres dimensiones.

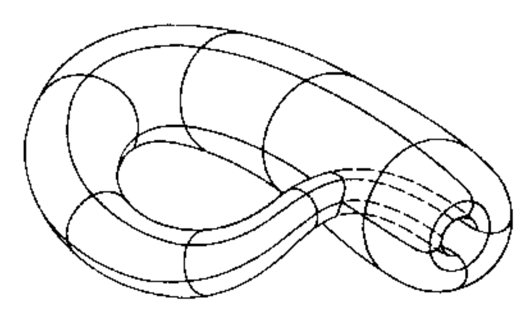


FIGURA 50.8

El plano proyectivo es otro ejemplo de una superficie cerrada bastante complicada; puede representarse topológicamente por un círculo con los pares de puntos diametralmente opuestos identifica-

dos (fig. 50.9), quedando representada la recta del infinito por la semicircunferencia CAD. Esta superficie es cerrada y su número de conexión es 1, o bien su género 0. Puede engendrarse también pegando el borde de un círculo a lo largo de una banda de Möbius (que justamente tiene un solo borde), aunque de nuevo la figura no puede ser construida en tres dimensiones sin que resulte que coincidan puntos que deberían ser distintos.

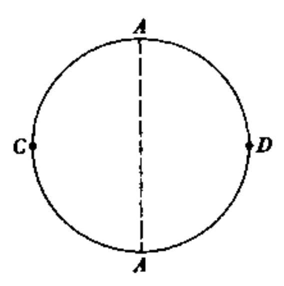


FIGURA 50.9

Aún otro impulso a la investigación topológica vino de la geometría algebraica. Ya hemos mencionado (cap. 39, sec. 8) que los geómetras se habían dedicado a estudiar las «superficies» de cuatro dimensiones que representan el dominio de las funciones algebraicas de dos variables complejas, y habían introducido integrales sobre estas superficies, de manera análoga a como se hace en la teoría de funciones e integrales algebraicas sobre las superficies de Riemann bidimensionales. Al estudiar estas figuras cuatridimensionales se investigó su tipo de conexión y se descubrió que no puede caracterizarse tales figuras por un número único, tal como el género caracteriza a las superficies de Riemann. Investigaciones llevadas a cabo por Emile Picard en torno a 1890 revelaron que para caracterizar dichas superficies serían necesarios por lo menos un número de conexión unidimensional y otro bidimensional.

La necesidad de estudiar el tipo de conexión de figuras de dimensión elevada fue reconocida por Enrico Betti (1823-1892), profesor de matemáticas en la universidad de Pisa, que decidió pasar directamente al caso de n dimensiones. Betti se había encontrado con Riemann en Italia, donde se había desplazado este último durante varios inviernos para cuidar su salud, y por él pudo tomar contacto

Betti con los trabajos de Riemann y de Clebsch. Betti [28] introdujo los números de conexión para cada dimensión de 1 a $n - 1$. El número de conexión unidimensional es el número de curvas cerradas que pueden dibujarse en la estructura geométrica y que no dividen a la superficie en regiones disjuntas (el número de conexión de Riemann era una unidad mayor). El número de conexión bidimensional es el número de superficies cerradas en la figura que, de una manera colectiva, no limiten ninguna región tridimensional de la misma, y de una manera análoga se definen los números de conexión para dimensiones más altas. Las curvas, superficies y figuras de dimensión mayor, cerradas, utilizadas en estas definiciones se llaman ciclos. (Si una superficie tiene borde, las curvas deben ser «secciones», es decir, curvas que van de un punto en el borde a otro punto del borde; por ejemplo, el número de conexión unidimensional de un cilindro hueco finito (sin extremos) es 1, porque puede trazarse una sección de un borde al otro sin desconectar la superficie). Betti demostró que el número de conexión unidimensional para las estructuras cuatridimensionales utilizadas para representar funciones algebraicas complejas $f(x,y,z) = 0$, es igual al número de conexión tridimensional.

4. La obra combinatoria de Poincaré

Hacia finales de siglo, el único campo que había sido cubierto de una manera bastante completa era la teoría de superficies cerradas; la obra de Betti era sólo el comienzo de una teoría más general. El autor que hizo el primer ataque sistemático y general a la teoría combinatoria de las figuras geométricas, y que suele ser considerado como el fundador de la topología, es Henri Poincaré (1854-1912). Poincaré fue profesor de matemáticas en la universidad de París, y se le conoce como el matemático más importante del último cuarto del siglo XIX y primeros años del siglo XX, así como el último hombre que tuvo un conocimiento universal de la matemática y de sus aplicaciones. Escribió un gran número de artículos de investigación, textos y artículos de divulgación, cubriendo casi todas las áreas básicas de la matemática y las más importantes de la física teórica, teoría del electromagnetismo, dinámica, mecánica de fluidos y astro-

---

[28] Annali di Mat., (2), 4, 1870-1871, 140-158.

nomía; su obra máxima es *Les Méthodes nouvelles de la mécanique céleste* (3 vols., 1892-1899). Los problemas físicos constituyeron la motivación de su investigación matemática.

Antes de ocuparse de la teoría combinatoria que vamos a describir, Poincaré había hecho contribuciones a otro campo de la topología, la teoría cualitativa de ecuaciones diferenciales (cap. 29, sec. 8); este trabajo era básicamente topológico porque se refería a la forma de las curvas integrales y a la naturaleza de los puntos singulares. Su contribución a la topología combinatoria estaba motivada por el problema de determinar la estructura de las superficies cuatridimensionales usadas para representar funciones algebraicas $f(x,y,z) = 0$, donde x, y, z son complejos. Decidió que era necesario un estudio sistemático del *análisis situs* de figuras generales n-dimensionales y, después de algunas notas en las «Comptes Rendus» de 1892 y 1893, publicó en 1895 [29] un trabajo básico, seguido por otros cinco largos suplementos esparcidos en varias revistas, hasta 1904. Poincaré consideraba su trabajo en topología combinatoria como una manera sistemática de estudiar la geometría n-dimensional, más bien que como un estudio de invariantes topológicos.

En su artículo de 1895, Poincaré intentaba abordar la teoría de figuras n-dimensionales usando sus representaciones analíticas, pero no hizo muchos progresos por este camino y volvió a una teoría puramente geométrica de variedades que son generalizaciones de las superficies de Riemann. Una figura es una variedad n-dimensional cerrada si cada punto posee entornos que son homeomorfos al interior n-dimensional de una $(n - 1)$-esfera; por ejemplo, la circunferencia (y cualquier figura homeomorfa) es una variedad unidimensional; una superficie esférica o un toro es una variedad bidimensional. Además de las variedades cerradas hay variedades con borde; un cubo y un toro sólido son variedades tridimensionales con borde; en un punto del borde un entorno es sólo parte del interior de una de dimensión 2.

El método que adoptó por último Poincaré aparece en su primer suplemento;[30] aunque él usaba celdas curvas o trozos de las figuras y variedades estudiadas, nosotros vamos a exponer sus ideas en términos de simplexes y complejos, que fueron introducidos posterior-

[29] Jour. de l'Ecole Poly., (2), 1, 1895, 1-121 = *Œuvres*, 6, 193-288.

[30] Rendiconti del Circolo Matematico di Palermo, 13, 1899, 285-343 = *Œuvres*, 6, 290-337.

mente por Brouwer. Un simplex es llanamente un triángulo n-dimensional; es decir, un simplex cero-dimensional es un punto, un simplex unidimensional es un segmento, un simplex bidimensional es un triángulo, uno tridimensional es un tetraedro, y un simplex n-dimensional es un tetraedro generalizado con $n + 1$ vértices. Las caras de dimensión menor de un simplex son a su vez simplexes. Un complejo es cualquier conjunto finito de simplex tales que la intersección de dos cualesquiera de ellos es una cara común, como máximo, y que toda cara de cada uno de los simplex del complejo es a su vez un simplex del complejo; a los simplex se les llama también celdas.

Para los fines de la teoría combinatoria se da una orientación determinada a cada simplex o celda de cualquier dimensión; por ejemplo, un 2-simplex E^2 (un triángulo) con vértices a_0, a_1 y a_2, puede ser orientado eligiendo un orden, digamos $a_0a_1a_2$, y entonces cualquier otro orden de los a_i que se obtenga a partir de éste por un número par de permutaciones de los a_i se dice que tiene la misma orientación. Así, E^2 viene dado por $(a_0a_1a_2)$ o $(a_2a_0a_1)$ o $(a_1a_2a_0)$. Cualquier orden derivable del original por un número impar de permutaciones de los a_i representa el simplex orientado de manera opuesta; por ejemplo, $-E^2$ viene dado por $(a_0a_2a_1)$ o $(a_1a_0a_2)$ o $(a_2a_1a_0)$.

El borde de un simplex está formado por los simplex de dimensión una unidad inferior contenidos en el dado, por ejemplo, el borde de un 2-simplex consiste en tres 1-simplex. Sin embargo, el borde debe ser tomado con la orientación correcta; así, para obtener el borde orientado se adopta la regla siguiente: el simplex E^k

$$a_0a_1a_2 \dots a_k$$

induce la orientación

$$(-1)^i(a_0a_1 \dots a_{i-1}a_{i+1} \dots a_k) \tag{1}$$

sobre cada uno de los simplex $(k-1)$-dimensionales que forman el borde. El E_i^{k-1} puede tener la orientación dada por (1), en cuyo caso el número de incidencia que representa la orientación de E_i^{k-1} respecto de E^k, es 1, o puede tener la orientación opuesta, en cuyo caso su número de incidencia es -1. Sea 1 o -1 el número de incidencia, el hecho básico es que el borde del borde de E^k es 0.

Dado un complejo cualquiera se puede formar combinaciones

lineales de sus simplex orientados k-dimensionales; por ejemplo, si los E_i^k son simplex k-dimensionales orientados, se tiene que

$$C^k = c_1E_1^k + c_2E_2^k + c_lE_l^k \tag{2}$$

donde los c_i son enteros positivos o negativos, es tal combinación lineal y se llama una cadena. Los números c_i nos dicen simplemente cuántas veces está contado cada simplex, significando el hecho de ser negativo que la orientación del simplex en cuestión está cambiada. Por ejemplo, si nuestra figura es un tetraedro determinado por cuatro puntos $(a_0a_1a_2a_3)$, podríamos formar la cadena $C^3 = 5\ E_1^3$. El borde de una cadena es la suma de todos los simplex de dimensión inmediatamente inferior de todos los simplex de la cadena, tomado cada uno de ellos con el número de incidencia correspondiente y con la multiplicidad que aparece en (2). Debido a que el borde de una cadena es la suma de los bordes de cada uno de los simplex que aparecen en ella, el borde del borde de una cadena es cero.

Una cadena cuyo borde es cero se llama un ciclo; es decir, algunas cadenas son ciclos. Entre los ciclos, algunos son bordes de otras cadenas, tal como por ejemplo el borde del simplex E_1^3 es un ciclo que limita a E_1^3. Sin embargo, si nuestra figura inicial no fuera el simplex tridimensional sino sólo su superficie, aún tendríamos el mismo ciclo, pero no sería borde en la figura considerada, es decir, en la superficie. Otro ejemplo: la cadena (fig. 50.10).

$$C_1^1 = (a_0a_1) + (a_1a_2) + (a_2a_3) + \dots + (a_5a_0)$$

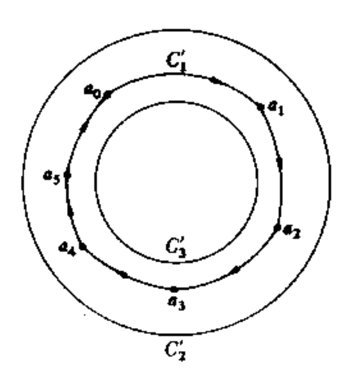


FIGURA 50.10

es un ciclo porque el borde de a_1a_2, por ejemplo, es $a_2 - a_1$ y el borde de la cadena completa es cero; sin embargo C_1 no es borde de ninguna cadena bidimensional. Esto es intuitivamente evidente porque el complejo en cuestión es un anillo circular y el agujero interior no forma parte de la figura.

Es posible que dos ciclos por separado no limiten ninguna región pero sí su suma o diferencia; por ejemplo, la suma (o diferencia) de C_2^1 y de C_3^1 (fig. 50.10) limita el área correspondiente al anillo completo; dos ciclos como estos se llaman dependientes. En general, los ciclos $C_1^k, \ldots, C_r^k$ se llaman dependientes si

$$\sum_{i=1}^{r} c_i C_i^k$$

es borde y no todas las c_1 son cero.

Poincaré introdujo además los importantes parámetros que denominó números de Betti (en honor de Enrico Betti). Para cada dimensión de los posibles simplex de un complejo, el número de ciclos independientes de esa dimensión recibe el nombre de número de Betti correspondiente a esa dimensión (en realidad Poincaré usaba un número que es una unidad mayor que el número de conexión de Betti). Así, por ejemplo, en el caso del anillo el número de Betti cero-dimensional es 1 porque todo punto es un ciclo pero dos puntos limitan la sucesión de segmentos lineales que los unen; el número de Betti unidimensional es 1 porque hay 1-ciclos que no son bordes pero dos cualesquiera de ellos (su suma o diferencia) son borde; el número de Betti bidimensional del anillo es cero porque ninguna cadena de 2-simplex es un ciclo. Para darse cuenta de lo que significan estos números, uno puede comparar el anillo con el círculo mismo (incluido su interior); en esta última figura todo ciclo unidimensional es borde, de manera que el número de Betti unidimensional es cero.

También introdujo Poincaré en su artículo de 1899 lo que llamaba coeficientes de torsión. Es posible en estructuras más complicadas, por ejemplo en el plano proyectivo, tener un ciclo que no es borde pero tal que dos veces dicho ciclo lo es; por ejemplo, si los simplex están orientados tal como muestra la figura 50.11, el borde de los cuatro triángulos es dos veces la recta *BB* (debe tenerse en cuenta que *AB* y *BA* son el mismo segmento lineal); el número 2 recibe el nombre de coeficiente de torsión, y el ciclo correspondiente

BB un ciclo de torsión; puede haber un número finito de tales ciclos de torsión independientes.

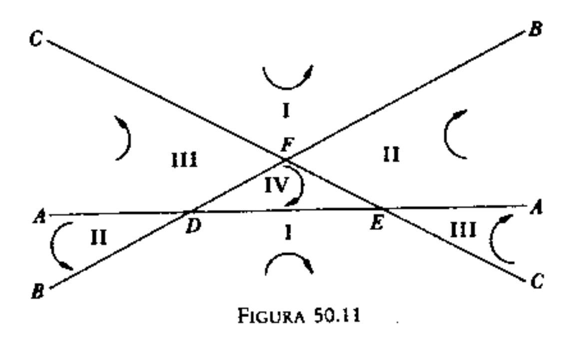


FIGURA 50.11

Incluso en estos ejemplos tan sencillos es claro que los números de Betti y los coeficientes de torsión de una figura geométrica distinguen de alguna manera una figura de otra, tal como se distingue el anillo circular del círculo.

En su primer suplemento (1899) y en el segundo,[31] Poincaré introdujo un método para calcular los números de Betti de un complejo. Cada simplex E_q de dimensión q tiene simplex de dimensión $q - 1$ en su borde, los cuales tienen números de incidencia $+1$ o -1; a un simplex $(q - 1)$-dimensional que no forme parte del borde de E_q se le hace corresponder el número de incidencia cero. Es posible entonces construir una matriz rectangular formada por los números de incidencia ϵ^q_{ij} del j-ésimo simplex q-dimensional con respecto al i-ésimo simplex q-dimensional. Existe una matriz T_q para cada dimensión q distinta de cero; T_q tendrá tantas filas como simplex q-dimensionales haya en el complejo, y tantas columnas como simplex $(q - 1)$-dimensionales. De esta forma, T_1 da las relaciones de incidencia de los vértices con las aristas, T_2 da las relaciones de incidencia de los simplex unidimensionales con los bidimensionales, y así sucesivamente. Mediante operaciones elementales sobre las matrices es posible hacer cero todos los elementos que no estén en la diagonal principal, y números enteros positivos o nulos los elementos de dicha diagonal. Supongamos que γ_q de estos elementos dia-

[31] Proc. London Math. Soc., 32, 1900, 277-308 = *Œuvres*, 6, 338-370.

gonales son 1; entonces demostró Poincaré que el número de Betti q-dimensional p_q (que es una unidad mayor que el número de conexión de Betti) es

$$p_q = \alpha_q - \gamma_{q+1} - \gamma_q + 1$$

donde α_q es el número de simplex q-dimensionales.

$$N(K^n) = \sum_{k=0}^{n} (-1)^k \alpha_k$$

Poincaré distinguía entre complejo con torsión y sin torsión; en el segundo caso todos los números que forman la diagonal principal son 0 o 1 para todo q, mientras que valores mayores indican la presencia de torsión.

También introdujo la característica $N(K^n)$ de un complejo n-dimensional K^n: si dicho complex tiene α_k simplex k-dimensionales, entonces, por definición

$$N(K^n) = \sum_{k=0}^{n} (-1)^k p_k$$

Este número es una generalización del número de Euler $V - A + C$. Poincaré demostró que si p_k es el k-ésimo número de Betti de K^n, entonces [32]

$$\sum_{i=1}^{n+1} x_i^2 = 1$$

resultado conocido como fórmula de Euler-Poincaré.

En su artículo de 1895 presenta Poincaré un teorema básico, con el nombre de teorema de dualidad, que se refiere a los números de Betti de una variedad cerrada. Una variedad cerrada n-dimensional, como dijimos anteriormente, es un complejo tal que cada uno de sus puntos tiene un entorno homeomorfo a una región del espacio euclídeo n-dimensional. El teorema afirma que en una variedad n-dimensional cerrada orientable, el número de Betti de dimensión p es

---

[32] Estos p_k son inferiores en una unidad a los de Poincaré. Aquí utilizamos el enunciado que es más familiar hoy.

igual al número de Betti de dimensión $n - p$; la demostración dada por Poincaré no era del todo completa, sin embargo.

En sus esfuerzos por distinguir unos complejos de otros introdujo Poincaré en 1895 otro concepto que juega actualmente un papel de una importancia considerable en topología, el grupo fundamental de un complejo, conocido también como el grupo de Poincaré o primer grupo de homotopía. La idea surge de considerar la diferencia entre regiones del plano simplemente y múltiplemente conexas; en el interior de un círculo todas las curvas cerradas pueden contraerse a un punto, pero en un anillo circular algunas de ellas, las que dejan en su interior la frontera circular más pequeña, no pueden ser contraídas a un punto, mientras que las curvas cerradas que no rodean dicha frontera interior sí pueden contraerse a un punto.

La mejor manera de introducir la idea precisa es considerar las curvas cerradas que empiezan y terminan en un punto dado y_0 del complejo; entonces, todas aquellas curvas que puedan ser deformadas una en otra por un movimiento continuo dentro del complejo, se llaman homótopas entre sí y son consideradas como una clase. Así, las curvas cerradas que comienzan y terminan en y_0 en el anillo circular (fig. 50.12) y que no encierran la frontera interior, forman una clase; las que comienzan y terminan en y_0 y encierran la frontera interior forman otra clase; las que comienzan y terminan en y_0 y rodean n veces dicha frontera constituyen otra clase.

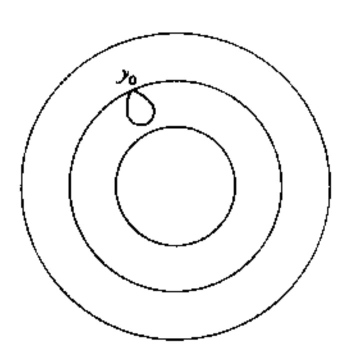


FIGURA 50.12

Entonces es posible definir una operación entre clases que geométricamente consiste en partir de y_0 y recorrer una curva cualquiera de la primera clase, y a continuación una curva cualquiera de la

segunda clase. El orden en que se toman las dos curvas así como el sentido de recorrido sobre cada una de ellas es importante. Las clases forman entonces un grupo no abeliano llamado el grupo fundamental del complejo con respecto al punto base y_0, y representado por $\pi_1(K, y_0)$ siendo K el complejo. Para complejos razonablemente sencillos el grupo no depende en realidad del punto base, es decir, los grupos correspondientes a y_0 e y_1, digamos, son isomorfos. Para el anillo circular el grupo fundamental es un grupo cíclico infinito, mientras que el dominio simplemente conexo que utiliza normalmente el análisis, es decir, el círculo cerrado, tiene un grupo fundamental trivial reducido al elemento identidad. De la misma forma que el círculo y el anillo circular se diferencian en sus grupos de homotopía, es posible describir complejos de mayor número de dimensiones que difieren entre sí marcadamente en este respecto.

Poincaré nos dejó algunas conjeturas importantes; por ejemplo, en su segundo suplemento afirma que dos variedades cerradas cualesquiera que tengan los mismos números de Betti y los mismos coeficientes de torsión, deben ser homeomorfas; sin embargo, en el quinto suplemento [33] dio un ejemplo de una variedad tridimensional que tenía los mismos números de Betti y coeficientes de torsión que la esfera tridimensional (la superfice de una esfera sólida cuatridimensional) pero que no era simplemente conexa. En vista de ello, añadió la conexión simple como una condición más, demostrando que hay variedades tridimensionales con los mismos números de Betti y coeficientes de torsión, pero que tienen grupos fundamenta-
y grupo fundamental sin ser por ello homeomorfas. Sin embargo, James W. Alexander (1888-1971), profesor de matemáticas en la universidad de Princeton y después en el Institute of Advanced Study, demostró [34] que dos variedades de dimensión tres pueden tener los mismos números de Betti, coeficientes de torsión y grupo fundamental sin ser por ello homeomorfas.

Sin embargo, James W. Alexander (1888-1971), profesor de matemáticas en la universidad de Princeton y después en el Institute of Advanced Study, demostró que dos variedades de dimensión tres pueden tener los mismos números de Betti, coeficientes de torsión y grupo fundamental sin ser por ello homeomorfas.

---

[33] Rendiconti del Circolo Matematico di Palermo, 18, 1904, 45-110 = *Œuvres*, 6, 435-498.

[34] Amer. Math. Soc. Trans., 20, 1919, 339-342.

En su quinto suplemento (1904) hizo Poincaré una conjetura un poco más restringida, consistente en que toda variedad tridimensional cerrada, orientable y simplemente conexa es homeomorfa a la esfera de la misma dimensión. Esta famosa conjetura ha sido generalizada en la forma: toda variedad n-dimensional cerrada y simplemente conexa que tenga los mismos números de Betti y los mismos coeficientes de torsión que la esfera n-dimensional, es homeomorfa a ella; ni la conjetura de Poincaré ni la generalizada ha sido demostradas (**).[35]

Otra famosa conjetura de Poincaré, llamada la «*Hauptvermutung*» (conjetura principal), afirma que si T_1 y T_2 son subdivisiones simpliciales de una misma 3-variedad, entonces T_1 y T_2 tienen subdivisiones isomorfas.[36]

5. Los invariantes combinatorios

El problema de establecer la invariancia de ciertas propiedades combinatorias consiste en demostrar que cualquier complejo homeomorfo a uno dado, considerados ambos como conjuntos de puntos, tiene las mismas propiedades combinatorias que el complejo dado. La demostración de que los números de Betti y los coeficientes de torsión son invariantes combinatorios fue dada por vez primera por Alexander,[37] probando que si K y K_1 son subdivisiones simpliciales cualesquiera (no necesariamente rectilíneas) de dos poliedros homeomorfos (como conjuntos de puntos) P y P_1, entonces los números de Betti y los coeficientes de torsión de P y P_1 son iguales. El recíproco no es cierto, de manera que la igualdad de los números

(**) La conjetura de Poincaré ha sido resuelta para $n = 4$ por M. Freedman. Véase M. Freedman, The topology of four dimensional manifolds, J. Diff. Geom., 17 (1983), 357-454. (Nota del traductor.)

[35] La conjetura generalizada ha sido demostrada para $n \geqslant 5$ por Stephen Smale (Amer. Math. Soc. Bull., 66, 1960, 373-375), John R. Stallings (*ibid.*, 485-488) y E. C. Zeeman (*ibid.*, 67, 1961, 270).

[36] La *Hauptvermutung* se ha demostrado para complejos simpliciales finitos (que son más generales que las variedades) de dimensión menor que 3, y es falsa para los de dimensión mayor que 5. Es correcta para variedades de dimensión menor o igual que 3 y está abierta para variedades de dimensión igual o mayor que 4. Ver John Milnor, Annals of Math., (2), 74, 1961, 575-590.

[37] Amer. Math. Soc. Trans., 16, 1915, 148-154.

de Betti y de los coeficientes de torsión de dos complejos no garantiza que dichos complejos sean homeomorfos.

Otro invariante de importancia se debe a L. E. J. Brouwer, que se vio interesado en topología a partir del estudio de problemas en teoría de funciones. Brouwer intentaba demostrar que hay $3g - 3$ clases de superficies de Riemann conformemente equivalentes, de género $g > 1$, y este problema le condujo a considerar otros problemas topológicos relacionados con él; Brouwer demostró [38] la invariancia de la dimensión de un complejo, en el sentido siguiente: si K es una subdivisión simplicial n-dimensional de un poliedro P, entonces toda subdivisión simplicial de P y toda división de ese tipo de cualquier poliedro homeomorfo a P es también un complejo n-dimensional.

Las demostraciones de este teorema, así como las del de Alexander, hacen uso de un método debido a Brouwer,[39] llamado de aproximaciones simpliciales de las transformaciones continuas. Las transformaciones simpliciales (de simplex en simplex) no son otra cosa que las análogas de dimensión superior de las transformaciones continuas, mientras que las aproximaciones simpliciales de las transformaciones continuas son análogas a la aproximación lineal de las funciones continuas; si el dominio en el que se realiza la aproximación es pequeño, entonces la aproximación sirve para representar la transformación continua a efectos de demostraciones de invariancia.

6. Los teoremas de punto fijo

Aparte de servir para distinguir unos complejos de otros, los métodos combinatorios han permitido obtener los llamados teoremas del punto fijo, que son geométricamente importantes y tienen también aplicaciones en análisis. Brouwer fue capaz, mediante la introducción de conceptos (en los que no entraremos) tales como la clase de una aplicación de un complejo en otro [40] y el grado de una aplicación,[41] de tratar lo que se llaman puntos singulares de los cam-

---

[38] Math. Ann., 70, 1910-1911, 161-165 y 71, 1911-1912, 305-313.

[39] Math. Ann., 71, 1911-1912, 97-115.

[40] Proceedings Koninklijke Akademie von Wettenschappen te Amsterdam, 12, 1910, 785-794.

[41] Math. Ann., 71, 1911-1912, 97-115.

pos vectoriales sobre una variedad, en primer lugar. Consideremos la circunferencia S^1, la superficie S^2 y la n-esfera $\Sigma_{i=1}^{n+1} x_i^2 = 1$ en el espacio euclídeo $(n + 1)$-dimensional; sobre S^1 es posible tener un vector tangente en cada punto, de tal forma que las longitudes y las direcciones de estos vectores varían de manera continua a lo largo de la circunferencia y ningún vector tiene longitud cero; se dice entonces que se tiene un campo vectorial tangente continuo y sin singularidades sobre S^1. Sin embargo no puede existir un campo así sobre S^2, y Brouwer demostró [42] que lo que pasa en S^2 debe pasar en toda esfera de dimensión par, es decir, no pueden existir campos vectoriales continuos sobre una esfera de dimensión par, que no tengan al menos un punto singular.

Relacionada estrechamente con la teoría de puntos singulares está la teoría de transformaciones continuas de unos complejos en otros, y de interés especial en tales transformaciones son los puntos fijos. Si denotamos por $f(x)$ el transformado de un punto x, entonces un punto fijo es el que verifica $f(x) = x$; para todo punto x podemos considerar el vector que va de x a $f(x)$, y entonces en el caso de un punto fijo el vector sería cero y dicho punto un punto singular del campo vectorial resultante. El teorema fundamental sobre existencia de puntos fijos es debido a Brouwer,[43] se aplica a simplex n-dimensionales (o bien homeomorfos a ellos), y afirma que toda transformación continua de un n-simplex en sí mismo tiene al menos un punto fijo. Así, por ejemplo, una transformación continua de un disco circular en sí mismo, debe tener al menos un punto fijo. En el mismo artículo demostró Brouwer que cualquier transformación continua inyectiva de una esfera de dimensión par en sí misma, que se pueda deformar en la transformación identidad, debe tener por lo menos un punto fijo.

Poco antes de su muerte en 1912, Poincaré demostró [44] que existirían órbitas periódicas en un cierto problema restringido de los tres cuerpos, siempre que se verificase un cierto teorema topológico; dicho teorema afirma que existen al menos dos puntos fijos en una región anular contenida entre dos circunferencias, cuando se somete dicha región a una transformación topológica que transforma cada

---

[42] Math. Ann., 71, 1911-1912, 97-115.

[43] Math. Ann., 71, 1911-1912, 97-115.

[44] Rendiconti del Circolo Matematico di Palermo, 33, 1912, 375-407 = *Œuvres*, 6, 499-538.

circunferencia en sí misma moviendo una en una dirección y la otra en la dirección opuesta, y que conserva el área. Este «último teorema» de Poincaré fue demostrado por George D. Birkhoff.[45]

Los teoremas de punto fijo fueron generalizados a espacios funcionales de dimensión infinita por Birkhoff y Olivier D. Kellog en un trabajo conjunto,[46] y fueron aplicados por Jules P. Schauder (1899-1940) [47] y conjuntamente por Schauder y Jean Leray (1906),[48] para demostrar la existencia de soluciones de ecuaciones diferenciales. Un teorema clave en tales aplicaciones dice que si T es una aplicación continua de un conjunto cerrado compacto y convexo de un espacio de Banach en sí mismo, entonces T tiene un punto fijo.

El uso de teoremas del punto fijo para demostrar la existencia de soluciones de ecuaciones diferenciales se entenderá mejor mediante un ejemplo sencillo. Consideremos la ecuación diferencial

$$\frac{dy}{dx} = F(x, y)$$

en el intervalo $0 \leqslant x \leqslant 1$, y la condición inicial $y = 0$ en $x = 0$. La solución $\Phi(x)$ satisface la ecuación

$$\phi(x) = \int_0^x F(x, \phi(x))\, dx$$

Podemos entonces definir la transformación general

$$g(x) = \int_0^x F(x, f(x))dx$$

donde $f(x)$ es una función arbitraria; esta transformación asocia la función g a la f, y se puede demostrar que es continua sobre el espacio de las funciones continuas $f(x)$ definidas en $[0,1]$. La solución Φ que buscamos es un punto fijo de este espacio de funciones. Si se puede demostrar que este espacio funcional satisface las condicio-

---

45 Amer. Math. Soc. Trans., 14, 1913, 14-22 = *Coll. Math. Papers*, 1, 673-681.

46 Amer. Math. Soc. Trans., 23, 1922, 96-115 = Birkhoff, *Coll. Math. Papers*, 3, 255-274.

47 Studia Mathematica, 2, 1930, 170-179.

48 Ann. de l'Ecole Norm. Sup., 51, 1934, 45-78.

nes que garantizan la existencia de puntos fijos, entonces queda establecida la existencia de Φ; los teoremas de punto fijo aplicables a espacios funcionales garantizan eso precisamente. El método ilustrado por este ejemplo sencillo nos permite establecer la existencia de soluciones de las ecuaciones en derivadas parciales no lineales que aparecen usualmente en el cálculo de variaciones y en hidrodinámica.

7. Generalizaciones y extensiones

Las ideas de Poincaré y Brouwer fueron aprovechadas por un cierto número de matemáticos que ampliaron el campo de la topología hasta el punto de que hoy día constituye uno de los dominios más activos de la matemática. Brouwer mismo extendió el teorema de la curva de Jordan;[49] este teorema (cap. 42, sec. 5) puede formularse de la manera siguiente: sea S^2 una 2-esfera (superficie) y sea J una curva cerrada (topológicamente una S^1) en S^2, entonces el número de Betti cero-dimensional de $S^2 - J$ es 2. Ya que este número de Betti es el número de componentes, J separa S^2 en dos regiones. La generalización de Brouwer afirma que una variedad $(n-1)$-dimensional separa el espacio euclídeo n-dimensional R_n en dos regiones. Alexander [50] generalizó el teorema de dualidad de Poincaré e indirectamente el teorema de la curva de Jordan; el teorema de Alexander afirma que el número de Betti r-dimensional de un complejo K en la esfera n-dimensional S^n es igual al número de Betti $(n-r-1)$-dimensional del dominio complementario $S^n - K$, siendo $r \neq 0$ y $r \neq n-1$. Para $r = 0$ el número de Betti de K es igual a 1 más el número de Betti $(n-1)$-dimensional de $S^n - K$, y para $r = n - 1$ el número de Betti de K es igual al número de Betti cero-dimensional de $S^n - K$ menos 1.[51] Este teorema generaliza el teorema de la curva de Jordan, pues si tomamos como K la esfera $(n-1)$-dimensional S^{n-1} entonces el teorema asegura que el número de Betti $(n-1)$-dimensional de S^{n-1}, que es 1, es igual al número de Betti cero-dimensional de $S^n - S^{n-1}$ menos 1, de manera que el

[49] Math. Ann., 71, 1911-1912, 314-319.
[50] Amer. Math. Soc. Trans., 23, 1922, 333-349.
[51] Alexander formuló este teorema bajo la condición de que los coeficientes de sus cadenas fueran enteros módulo 2 (véase parágrafo siguiente). Nuestra formulación es para coeficientes enteros usuales.

número de Betti cero-dimensional de $S^n - S^{n-1}$ es 2 y S^{n-1} divide a S^n en dos regiones.

Las definiciones de los números de Betti han sido modificadas y generalizadas en varios sentidos. Veblen y Alexander [52] introdujeron las cadenas y ciclos módulo 2, es decir, en vez de simplex orientados se emplean simplex no orientados, pero los coeficientes enteros se toman módulo 2, y los bordes de cadenas se cuentan de la misma manera; posteriormente Alexander [53] introdujo los coeficientes módulo m para cadenas y ciclos. Solomon Lefschetz (1884-1973) sugirió utilizar números racionales como coeficientes,[54] y Lev L. Pontrjagin (1908-1960) [55] hizo una generalización aún más amplia tomando como coeficientes de las cadenas elementos de un grupo abeliano; este último concepto incluye como casos particulares las clases de coeficientes mencionados más arriba, así como otra clase también utilizada, la de los números reales módulo 1. Todas estas generalizaciones, aunque conducían a teoremas más generales, no mejoran la capacidad de los números de Betti y de los coeficientes de torsión para distinguir complejos.

Otro cambio en la formulación de las propiedades combinatorias básicas, que fue llevado a cabo durante los años 1925 a 1930 por diversos matemáticos, posiblemente por sugerencia de Emmy Noether, fue el de reformular la teoría de cadenas y ciclos en el lenguaje de la teoría de grupos. Cadenas de la misma dimensión pueden sumarse unas con otras de la manera natural, es decir, sumando los coeficientes del mismo simplex, y debido a que los ciclos son cadenas pueden también sumarse ciclos y su resultado es otro ciclo. De esta forma las cadenas y los ciclos forman grupos. Dado un complejo K, para cada cadena k-dimensional existe el borde $(k-1)$-dimensional, y la suma de dos cadenas tiene como borde la suma de los bordes de las cadenas individuales; por tanto, la relación de cadena a borde establece un homomorfismo del grupo $C^k(K)$ de las cadenas k-dimensionales en un subgrupo $H^{k-1}(K)$ del grupo de las cadenas $(k-1)$-dimensionales. El conjunto de todos los k-ciclos ($k>0$) es un subgrupo $Z^k(K)$ de $C^k(K)$ que se transforma por este homomorfismo en el elemento identidad o 0 de $C^{k-1}(K)$. Debido a que el borde de toda cadena es un ciclo, $H^{k-1}(K)$ es un subgrupo de $Z^{k-1}(K)$.

[52] Annals of Math., (2), 14, 1913, 163-178.
[53] Amer. Math. Soc. Trans., 28, 1926, 301-329.
[54] Annals of Math., (2), 29, 1928, 232-254.
[55] Annals of Math., (2), 35, 1934, 904-914.

Sobre la base de estos hechos podemos formular la definición siguiente: para todo $k \geqslant 0$, el grupo cociente de $Z^k(K)$, es decir, del grupo de los ciclos k-dimensionales, módulo el subgrupo $H^k(K)$ de los ciclos borde, recibe el nombre de grupo de homología k-ésimo de K, representado por $B^k(K)$. El número de generadores linealmente independientes de este grupo cociente, recibe el nombre de número de Betti k-ésimo del complejo, y se representa por $p^k(K)$. El grupo k-ésimo de homología puede contener también grupos cíclicos finitos, los cuales corresponden a los ciclos de torsión; de hecho, los órdenes de estos grupos finitos son los coeficientes de torsión. Con esta formulación de los grupos de homología de un complejo, muchos resultados antiguos pueden reformularse en el lenguaje de la teoría de grupos.

La generalización más importante de la primera parte de este siglo fue la de introducir la teoría de homología para espacios generales, tales como los espacios métricos compactos, en vez de partir de figuras que sean complejos. Las ideas básicas se deben a Paul S. Alexandroff (n. 1896),[56] Leopold Vietoris (1891)[57] y Eduard Cech (1893-1960).[58] No daremos aquí los detalles, ya que suponen una formulación totalmente nueva de la teoría de homología. Es necesario hacer observar, sin embargo, que este trabajo marca un hito en el proceso de fusión de la topología general o conjuntista y la topología combinatoria.

Bibliografía

Bouligand, Georges: *Les Définitions modernes de la dimension*, Hermann, 1935.

Dehn, M., y Heegard, P.: «Analysis Situs». *Encyk. der Math. Wiss.*, B. G. Teubner, 1907-1910, III AB3, 153-220.

Franklin, Philip: «The Four Color Problem». *Scripta Mathematica*, 6, 1939, 149-156, 197-210.

Hadamard, J.: «L'Œuvre mathématique de Poincaré». *Acta Math.*, 38, 1921, 203-287.

Manheim, J. H.: *The Genesis of Point Set Topology*, Pergamon Press, 1964.

---

[56] Annals of Math., (2), 30, 1928-1929, 101-187.

[57] Math. Ann., 97, 1927, 454-472.

[58] Fundamenta Mathematicae, 19, 1932, 149-183.

Osgood, William F.: «Topics in the Theory of Functions of Several Complex Variables». *Madison Colloquium*, American Mathematical Society, 1914, 111-230.

Poincaré, Henri: *Œuvres*, Gauthier-Villars, 1916-1956, vol. 6.

Smith, David E.: *A Source Book in Mathematics,* Dover (reimpresión), 1959, vol. 2, 404-410.

Tietze, H., y Vietoris L.: «Beziehungen zwischen den verschiedenen Zweigen der Topologie». *Encyk. der Math. Wiss.*, B. G. Teubner, 1914-1931, III AB13, 141-237.

Zoretti, L., y Rosenthal, A.: «Die Punktmengen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, II C9A, 855-1030.

Capítulo 51
LOS FUNDAMENTOS DE LA MATEMATICA

La lógica es invencible, porque para combatir la lógica es necesario utilizar la lógica misma.

PIERRE BOUTROUX

Sabemos que los matemáticos no se preocupan más por la lógica que los lógicos por la matemática. Los dos ojos de que dispone la ciencia exacta son la matemática y la lógica; la secta matemática extrae el ojo lógico, la secta lógica extrae el ojo matemático, pensando cada uno que puede ver mejor con un ojo que con dos.

AUGUSTUS DE MORGAN

1. Introducción

La actividad más profunda de la matemática del siglo XX, con gran diferencia sobre las demás, ha sido la investigación sobre sus propios fundamentos. Los problemas que le vinieron impuestos al respecto a los matemáticos, y otros que ellos asumieron voluntariamente, se refieren no sólo a la naturaleza misma de la matemática, sino también a la validez de la matemática deductiva.

Hubo varias circunstancias que convergieron para poner en primera línea los problemas de fundamentos a principios de siglo. La primera fue el descubrimiento de las contradicciones, llamadas eufemísticamente paradojas, especialmente en la teoría de conjuntos. De una de estas contradicciones, la paradoja de Burali-Forti, hemos hablado ya (cap. 41, sec. 9), y durante los primeros años de este siglo se descubrieron varias otras. Naturalmente, tales descubrimientos inquietaron profundamente a algunos matemáticos. Otro problema que había ido surgiendo gradualmente y que al fin apareció en escena en

sus verdaderas proporciones, fue el de la consistencia de la matemática (cap. 43, sec.6). En vista de las paradojas de la teoría de conjuntos, era allí donde urgía especialmente establecer la consistencia.

Durante los últimos años del siglo XIX hubo un cierto número de matemáticos que comenzaron a preocuparse de los fundamentos de su ciencia y, de una manera muy particular, de las relaciones entre la matemática y la lógica. Las investigaciones en este área (de cuyos detalles hablaremos más adelante) sugirieron a algunos matemáticos la idea de que la matemática podría fundamentarse sobre la lógica. Otros, en cambio, pusieron en cuestión la aplicación universal de los principios lógicos, el dudoso significado de algunas demostraciones existenciales, e incluso la fiabilidad de la demostración lógica como justificación de los resultados matemáticos obtenidos. Las controversias, que habían permanecido en estado latente como los rescoldos antes de 1900, se convirtieron en un fuego declarado cuando las paradojas y los problemas de consistencia añadieron combustible. Inmediatamente se volvió vital y de un interés general la cuestión de la fundamentación adecuada de toda la matemática.

2. Las paradojas de la teoría de conjuntos

Poco después del descubrimiento por Cantor y Burali-Forti de la paradoja sobre los números ordinales, aparecieron varias otras paradojas o antinomias. En realidad la palabra paradoja es ambigua e inadecuada en este contexto, porque parece referirse a una contradicción sólo aparente, mientras que lo que se habían encontrado los matemáticos eran incuestionables y palmarias contradicciones. Veamos primero en qué consistían.

Una de las paradojas fue popularizada por Bertrand Russell (1872-1970) en 1918, con el nombre de «paradoja del barbero». El barbero de un pueblecillo, presumiendo de no tener competencia, se anuncia diciendo que él no afeita a aquellos que se afeitan a sí mismos, desde luego, pero sí afeita a todos aquellos que no se afeitan a sí mismos. Un buen día se le ocurre preguntarse si debería afeitarse a sí mismo o no. Si se afeita a sí mismo, entonces, por la primera parte de su afirmación, no debería afeitarse a sí mismo. Pero si no se afeita a sí mismo, entonces por la segunda parte debería afeitarse a sí mismo. El pobre barbero se encuentra en un verdadero atolladero lógico.

Otra paradoja formulada por Jules Richard (n. 1862),[1] de la que G. G. Berry y Russell dieron una versión simplificada, fue publicada por el segundo de ellos.[2] La paradoja simplificada, conocida también como «paradoja de Richard», dice así: todo número natural puede definirse utilizando palabras con un cierto número de letras, en general de diversas maneras. Por ejemplo, el número 36 puede definirse como «treinta y seis» o como «cuatro por nueve»; la primera definición contiene doce letras y la segunda catorce. No hay una manera uniforme de definir cualquier número, pero esto no es esencial. Clasifiquemos ahora todos los números naturales en dos grupos; el primero incluirá todos aquellos que puedan definirse (de una manera por lo menos) con 100 letras o menos, mientras que el segundo incluirá todos aquellos que, de cualquier manera que se definan, se necesiten para ello al menos 101 letras. Es obvio que sólo hay un número finito de números que se puedan definir con 100 letras o menos, puesto que hay exactamente 27^{100} expresiones con 100 letras (y la mayoría no tienen sentido). Existe por tanto el número mínimo del segundo grupo que, evidentemente, puede definirse por la frase «el mínimo número natural que no se puede definir con cien letras o menos». Pero esta frase tiene menos de 100 letras. Luego el mínimo número natural no definible con 100 letras o menos puede definirse con menos de 100 letras.

Otra forma de esta paradoja fue enunciada por primera vez en 1908 por Kurt Grelling (1886-1941) y Leonard Nelson (1882-1927), que la publicaron en una revista muy poco conocida.[3] Algunas palabras (adjetivos concretamente) se describen o aplican a sí mismas y otras no. Por ejemplo, la palabra «polisílaba» es polisílaba; «corta» es corta. En cambio la palabra «monosílaba» no es monosílaba y «larga» no es larga. Llamemos a todas aquellas que se aplican a sí mismas «autológicas», y a las que no se aplican a sí mismas «heterológicas». Así pues, «*X*» es heterológica si «*X*» no tiene la propiedad *X*. Si lo aplicamos ahora a la palabra heterológica, sustituyéndola por la *X*, nos queda que «heterológica» es heterológica si «heterológica» no es heterológica.

Cantor señalaba, en carta a Dedekind de 1899, que no se podría hablar del conjunto de todos los conjuntos sin entrar en contradic-

---

[1] Revue Générale des Sciences, 16, 1905, 541.
[2] Proc. London Math. Soc., (2), 4, 1906, 29-53.
[3] *Abhandlungen der Friesschen Schule*, 2, 1908, 301-324.

ción (cap. 41, sec. 9). Algo análogo ocurre con la paradoja de Russell (*The Principles of Mathematics*, 1903, p. 101). La clase de todos los hombres no es un hombre, pero la clase de todas las ideas es una idea; la clase de todas las bibliotecas es una biblioteca, y la clase de todos los conjuntos de cardinal mayor que uno tiene cardinal mayor que uno. Por tanto, algunas clases no son elementos de sí mismas y otras sí lo son. Esta clasificación las abarca a todas y los dos tipos son mutuamente excluyentes. Sea M la clase de todas las clases que son elementos de sí mismas, y N la clase de todas las clases que no son elementos de sí mismas. Podemos preguntarnos ahora si la clase N cae dentro de M o de N. Si N perteneciese a N entonces sería elemento de sí misma y por tanto debería pertenecer a M. Por otro lado, si N pertenece a M, entonces, dado que M y N son mutuamente excluyentes, N no podría pertenecer a N. Por tanto, N no es elemento de sí misma y, en virtud de su propia definición, debería pertenecer a N.

La causa de todas estas paradojas, como señalan Russell y Whitehead, radica en la definición de un objeto en términos de una clase que contiene como elemento al objeto que se está definiendo. Tales definiciones se llaman impredicativas y aparecen de manera especial en teoría de conjuntos. Como hizo observar Zermelo en 1908, este tipo de definición se utiliza también para definir el extremo superior de un conjunto acotado de números reales y otros conceptos del análisis. Así pues, el análisis contiene paradojas.

La demostración de Cantor de la no numerabilidad del conjunto de los números reales (cap. 41, sec. 7) también utiliza un conjunto definido impredicativamente. Se supone dada una correspondencia biunívoca f entre el conjunto de todos los números naturales y el conjunto M de todos los números reales (identificados a conjuntos de números naturales). A cada número natural k corresponde el conjunto $f(k)$. Ahora bien, k pertenece o no pertenece a $f(k)$. Sea N el conjunto de todos los k tales que k no pertenece a $f(k)$; este conjunto N es un número real y, por tanto, por la correspondencia biunívoca f tiene que haber un único natural n tal que $f(n) = N$. Ahora bien, si n pertenece a N tendría que no pertenecer, por la definición misma de N; pero si no perteneciese debería pertenecer por la misma razón. No puede existir, por tanto, la correspondencia f. La definición del conjunto N es impredicativa porque k pertenece a N si y sólo si existe un conjunto K en M tal que $K = f(k)$ y k no pertenece a K. Así pues, al definir N hacemos uso de la totalidad

de los conjuntos *M* que contienen a *N* como elemento; por tanto, para definir *N*, *N* ha de estar ya en *M*.

Es muy fácil caer inadvertidamente en la trampa de introducir definiciones impredicativas. Así, si uno define la clase de todas las clases que contiene más de cinco elementos, ha definido una clase que se contiene a sí misma como elemento. Análogamente, la frase «el conjunto *S* de todos los conjuntos definibles con treinta palabras o menos», define a *S* impredicativamente.

Estas paradojas inquietaron a muchos matemáticos, dado que ponían en cuestión no sólo la teoría de conjuntos, sino también grandes dominios del análisis clásico. La matemática, como estructura lógica, se encontraba en un estado bien triste, y los matemáticos pensaban con añoranza en los viejos y felices días de antes de que aparecieran las paradojas.

3. La axiomatización de la teoría de conjuntos

No nos puede sorprender el hecho de que una de las primeras medidas que tomaran los matemáticos fuera la de axiomatizar la teoría de conjuntos que había formulado Cantor de una manera muy libre o informal o, como algunos prefieren decir hoy, intuitiva. La axiomatización de la geometría y de la aritmética había permitido resolver problemas lógicos en esas ramas, y parecía verosímil que la axiomatización también clarificaría las dificultades de la teoría de conjuntos. El primero en emprender esta tarea fue el matemático alemán Ernst Zermelo, que creía que las paradojas habían aparecido porque Cantor no había restringido adecuadamente el concepto de conjunto. En 1895 [4] Cantor había definido un conjunto como una colección de objetos distintos de nuestra intuición o nuestro pensamiento. Esto era bastante vago, y por tanto Zermelo esperaba que un sistema de axiomas precisos y explícitos clarificaría qué es lo que se entiende por un conjunto y cuáles son sus propiedades. Cantor mismo no había dejado de ser consciente de que su concepto de conjunto presentaba dificultades; en una carta a Dedekind de 1899 [5] distinguía entre conjuntos consistentes e inconsistentes. Zermelo pensó que podría restringir sus conjuntos a los consistentes de Cantor,

---

[4] Math. Ann., 46, 1895, 481-512 = *Ges. Abh.*, 282-356.
[5] *Ges. Abh.*, 443-448.

y que éstos serían suficientes para la matemática. Su sistema de axiomas [6] contenía conceptos y relaciones fundamentales que estaban definidas implícitamente por las afirmaciones de los axiomas mismos. Entre tales conceptos los fundamentales eran el de conjunto y la relación de pertenencia de un elemento a un conjunto. No debería utilizarse ninguna propiedad de los conjuntos salvo que estuviese garantizada por un axioma. La existencia de un conjuno infinito y las operaciones tales como la unión de conjuntos o la formación de subconjuntos también estaban especificadas en dichos axiomas. Zermelo incluyó de manera destacada el llamado «axioma de elección» (cap. 41, sec. 8).

El plan de Zermelo era el de admitir en la teoría de conjuntos sólo aquellas clases de las que verosímilmente no pudieran derivarse contradicciones. Así, por ejemplo, la clase vacía, cualquier clase finita y la clase de los números naturales parecían seguras. Dada una clase segura, ciertas otras clases formadas a partir de ella, tales como cualquiera de sus subclases, la clase de todas sus subclases y la unión de clases seguras, deberían ser a su vez clases seguras. Evitó, sin embargo, la complementación, puesto que, aunque x sea una clase segura, el complemento de x, es decir, todos los no-x en algún universo de objetos muy grande podría no ser segura.

La fundamentación de la teoría de conjuntos de Zermelo fue mejorada por Abraham A. Fraenkel (1891-1965),[7] y von Neumann [8] introdujo cambios adicionales. En el caso del sistema de Zermelo-Fraenkel, la esperanza de evitar las paradojas se basa en restringir los tipos de conjuntos que se admiten, siempre que los admitidos sean suficientes para las necesidades del análisis. La idea de von Neumann era un poco más atrevida, estableciendo una distinción entre conjuntos y clases propias. Las clases propias eran colecciones tan grandes que no estaban contenidas como elementos en ninguna otra clase o conjunto, mientras que los conjuntos eran colecciones más restringidas, que podían ser elementos de otras clases. Así pues, los conjuntos eran exactamente las clases seguras y, como apuntaba von Neumann, no era la admisión de las clases propias lo que conducía a contradicciones, sino el tratarlas como elementos de otras clases.

---

[6] Math. Ann., 65, 1908, 261-281.

[7] Math. Ann., 86, 1921-1922, 230-237, y muchos artículos posteriores.

[8] Jour. für Math., 154, 1925, 219-240 y artículos posteriores.

La teoría de conjuntos formal de Zermelo, modificada por Fraenkel, von Neumann y otros, resulta adecuada para desarrollar la teoría de conjuntos que se necesita para todo el análisis clásico prácticamente, y evitar las paradojas en el sentido de que, hasta hoy, nadie ha descubierto ninguna dentro de la teoría. Sin embargo, la consistencia de la teoría axiomática de conjuntos no ha sido demostrada. A propósito de esta cuestión abierta de la consistencia, decía Poincaré: «Hemos puesto una valla en torno al rebaño para protegerlo de los lobos, pero lo que no sabemos es si habrá quedado algún lobo dentro ya de la valla.»

Aparte del problema de la consistencia, hay que decir que la axiomatización de la teoría de conjuntos incluía el axioma de elección, necesario para establecer importantes resultados del análisis clásico, de la topología y del álgebra abstracta. Un cierto número de matemáticos, entre los que se contaban Hadamard, Lebesgue, Borel y Baire, consideraron objetable este axioma, y en 1904, cuando Zermelo lo utilizó para demostrar el teorema de buena ordenación (cap. 41,sec. 8), una verdadera avalancha de objeciones inundó las revistas matemáticas.[9] Las cuestiones acerca de si este axioma era esencial para los resultados en los que se utilizaba, y de si era independiente de los restantes o no, no tardaron en plantearse, permaneciendo durante muchos años sin ser resueltas (véase la sec. 8).

La axiomatización de la teoría de conjuntos, a pesar de que dejaba abiertas cuestiones como el problema de la consistencia y el papel del axioma de elección, podría haber tranquilizado al menos a los matemáticos con respecto a las paradojas, haciendo declinar así el interés por los fundamentos. Pero para entonces ya habían entrado en actividad y en plena polémica varias escuelas filosóficas sobre los fundamentos de la matemática, motivadas sin duda por la paradojas y el problema de la consistencia. Los partidarios de esas líneas filosóficas no encontraban satisfactorio el método axiomático tal como lo practicaban Zermelo y otros. A algunos de ellos les resultaba cuestionable porque presuponía la lógica que utilizaba, mientras que por entonces ya estaban sometidas a investigación tanto la lógica misma como sus relaciones con la matemática. Otros, más radicales,

---

[9] Las opiniones de estos matemáticos aparecen en un famoso intercambio epistolar. Véase Bull. Soc. Math. de France, 33, 1905, 261-273. También en E. Borel, *Leçons sur la théorie des fonctions*, Gauthier-Villars, 4.ª ed., 1950, 150-158.

rechazaban apoyarse en cualquier tipo de lógica, especialmente al aplicarla a conjuntos infinitos. Para entender los argumentos que proponían las diversas escuelas de pensamiento, necesitamos retroceder un poco en el tiempo.

4. La aparición de la lógica matemática

Un tema que provocó nuevas controversias e insatisfacciones con la axiomatización de la teoría de conjuntos es el que se refiere al papel de la lógica en la matemática, que provenía ya de la matematización de la lógica llevada a cabo durante el siglo XIX. Esta corriente de desarrollo tiene su propia historia.

La potencia del álgebra para representar simbólicamente e incluso mecanizar argumentos geométricos, había impresionado ya a Descartes y a Leibniz, entre otros (cap. 13., sec. 8), y ambos concibieron una ciencia mucho más general que el álgebra de los números. Se trataría de una ciencia abstracta o general de razonamiento, que operaría de un modo algo parecido al álgebra ordinaria, pero que sería aplicable al razonamiento en todos los campos. Tal como lo expresaba Leibniz en uno de sus escritos, «la matemática universal es, por así decirlo, la lógica de la imaginación», y debería tratar de «todo aquello que en el dominio de la imaginación sea susceptible de determinación exacta». Con una lógica tal se podría construir cualquier edificio del pensamiento, a partir de sus elementos más simples, hasta alcanzar estructuras más y más complicadas. El álgebra universal sería parte de la lógica, pero de una lógica en forma algebraica. Descartes comenzó modestamente, intentando construir un álgebra de la lógica, de la que nos queda un bosquejo incompleto.

Leibniz, persiguiendo la misma meta general que Descartes, se propuso un programa más ambicioso. A lo largo de toda su vida había prestado atención a la lógica, y desde muy joven se había visto fascinado por el proyecto del teólogo medieval Raimundo Lulio (1235-1315), que en su libro *Ars Magna et Ultima* presentaba un ingenuo método mecánico para producir nuevas ideas combinando otras ya existentes, que implicaba la idea de una ciencia universal de la lógica que sería aplicable a todo tipo de razonamientos. Leibniz se apartó de la lógica escolástica y de Lulio, pero quedó impresionado por la posibilidad de un cálculo muy general que le permitiría al hombre razonar en todos los campos mecánicamente y sin esfuer-

zo alguno. Leibniz nos dice de su plan para una lógica simbólica universal que tal ciencia, de la que el álgebra ordinaria sólo sería una pequeña parte, sólo estaría limitada por la necesidad de obedecer las leyes de la lógica formal. Podríamos llamarla, decía él, una «síntesis algebraico-lógica».

Esta ciencia general suministraría, ante todo, un lenguaje racional universal, adaptado a todo tipo de pensamiento. Los conceptos, una vez resueltos en otros primitivos, distintos y separados que no se solapen, podrían combinarse de una manera casi mecánica. Leibniz pensó también que sería necesario un simbolismo especial para evitar a la mente perderse. Aquí se puede ver claramente la influencia del simbolismo algebraico en su pensamiento. Lo que pretendía era un lenguaje simbólico capaz de expresar los pensamientos humanos sin ambigüedades, y de ayudar a realizar deducciones. Este lenguaje simbólico sería su *characteristica universalis*.

En 1666 escribió Leibniz su *De Arte Combinatoria*,[10] que contiene, entre otras cosas, sus primeras ideas sobre el sistema universal de razonamiento. Más tarde escribiría numerosos otros fragmentos sobre el mismo tema que no llegaron a publicarse nunca, pero que pueden verse recogidos en la edición de sus escritos filosóficos (ver nota 10). En su primer ensayo asociaba a cada concepto primitivo un número primo; así, cualquier concepto compuesto de varios otros primitivos venía representado por el producto de los primos correspondientes. Por ejemplo, si 3 representa «animal» y 7 «racional», 21 representaría «animal racional». A continuación intentó trasladar las reglas usuales de los silogismos a este esquema pero no tuvo éxito. También intentó en otra ocasión utilizar símbolos especiales en vez de números primos, donde, una vez más, las ideas complejas se representarían por combinaciones de símbolos. En realidad, Leibniz pensaba que el número de ideas primitivas sería pequeño, cosa que se demostró errónea. Tampoco fue suficiente con una sola operación básica, la conjunción, para construir todas las composiciones de ideas primitivas.

También comenzó a trabajar Leibniz en un álgebra de la lógica propiamente dicha. Directa e indirectamente disponía en su álgebra de conceptos que hoy describiríamos como adición lógica, multipli-

---

[10] Pub. 1690 = G. W. Leibniz: *Die philosophischen Schriften*, ed. por C. I. Gerhardt, 1875-1890, vol. 4, 27-102.

cación lógica, identidad, negación y la clase vacía. También llamó la atención sobre el interés del estudio de relaciones abstractas, tales como la inclusión, las correspondencias biunívocas o multívocas y las relaciones de equivalencia, reconociendo que algunas de estas relaciones tenían las propiedades de simetría y transitividad. Leibniz no completó esta obra; no pasó de las reglas de los silogismos que, como él mismo tuvo que admitir, no agotan toda la lógica que utiliza la matemática. Leibniz comunicó sus ideas a l'Hôpital y a otros, pero no le prestaron atención y sus escritos lógicos permanecieron sin publicarse hasta comienzos del siglo XX, de manera que su influencia directa fue muy pequeña. Durante el siglo XVIII y comienzos del XIX algunos hicieron intentos análogos a los de Leibniz, pero no llegaron más lejos.

Una etapa más eficaz, aunque menos ambiciosa, fue la llevada a cabo por Augustus de Morgan. De Morgan publicó su *Formal Logic* en 1847 y muchos artículos, algunos de los cuales aparecieron en las *Transactions of the Cambridge Philosophical Society*, en los que trataba de corregir defectos de la lógica aritotélica y mejorarla. En su *Formal Logic* añadía un nuevo principio a la lógica aristotélica. En esta última las premisas «Algunos *M* son *A*» y «Algunos *M* son *B*» no permiten extraer ninguna conclusión; de hecho, esta lógica afirma que el término medio *M* ha de aparecer de manera universal, es decir, deben aparecer «Todos los *M*». De Morgan sostiene, sin embargo, que de «La mayor parte de los *M* son *A*» y de «La mayor parte de los *M* son *B*» se sigue necesariamente que «Algunos *A* son *B*». De Morgan razona este hecho de una forma cuantitativa. Si hay m de los *M*, y a de los *M* son *A*, y b de los *M* son *B*, entonces al menos $(a + b - m)$ de los *A* son *B*. El interés de la observación de de Morgan está en que los términos pueden ser cuantificados y, como consecuencia pudo introducir muchas más formas válidas de silogismo. La cuantificación también eliminó un defecto de la lógica aristotélica: la conclusión «Algunos *A* son *B*», que en lógica aristotélica se sigue de la premisa «Todos los *A* son *B*», implicaría la existencia de los *A*, que no tienen por qué existir.

De Morgan inició también el estudio de la lógica de relaciones. La lógica aristotélica está dedicada básicamente a la relación «ser», y se limita a afirmar o negar esta relación. Como hizo observar de Morgan, esta lógica no podría demostrar que si un caballo es un animal, entonces una cola de caballo es una cola de un animal, y ciertamente no podría manejar una relación tal como la «*x* ama a

y». De Morgan introdujo un simbolismo adecuado al manejo de relaciones, pero no llegó demasiado lejos.

En el campo de la lógica simbólica, de Morgan es bien conocido sobre todo por las hoy llamadas leyes de Morgan. Tal como él las formuló,[11] el contrario de un agregado es el compuesto de los contrarios de los agregados; y el contrario de un compuesto es el agregado de los contrarios de los componentes. En simbolismo lógico estas dos leyes se escriben

$$1 - (x + y) = (1 - x)(1 - y)$$
$$1 - xy = (1 - x) + (1 - y).$$

La contribución del simbolismo a un álgebra de la lógica constituye la mayor aportación de George Boole (1815-1864), que fue en gran medida un autodidacta y llegó a ser profesor de matemáticas en el Kings College de Cork. Boole estaba profundamente convencido de que la simbolización del lenguaje rigorizaría la lógica. Su obra *Mathematical Analysis of Logic*, que curiosamente apareció el mismo día que la *Formal Logic* de de Morgan, y *An Investigation of the Laws of Thought* (1854), contienen sus ideas más importantes.

El planteamiento básico de Boole consistía en hacer recaer el énfasis en la lógica extensional, es decir, en una lógica de clases, donde, por cierto, los conjuntos o clases se representaban por x, y, z, ..., mientras que los símbolos X, Y, Z, ... representaban a los miembros individuales. La clase universal venía representada por 1 y la clase vacía o nula por 0. Boole utilizó el término xy para representar la intersección de dos conjuntos x e y (operación a la que llamó elección), es decir, el conjunto de todos los elementos comunes a x y *a* y, y el $x + y$ para representar el conjunto de los elementos tanto de x como de y (para Boole, hablando estrictamente, la unión o adición sólo se aplicaba a conjuntos disjuntos; W. S. Jevons (1835-1882) generalizó este concepto). El complemento de x viene representado por $1 - x$ y, de una manera más general, $x - y$ es la clase de los elementos de x que no están en y. La relación de inclusión, es decir, x está contenido en y, la representó por $xy = x$, y en todos los casos el signo igual representa la identidad de dos clases.

Boole creía que nuestra mente nos otorga de modo inmediato ciertos procesos de razonamiento elementales que son los axiomas

---

[11] Trans. Camb. Phil. Soc., 10, 1858, 173-230.

de la lógica. Por ejemplo, el principio de contradicción, que *A* no puede ser nunca *B* y no *B* a la vez, es un hecho axiomático, que se expresa por la fórmula

$$x(1 - x) = 0$$

También resulta obvio para la mente que

$$xy = yx$$

y así, esta propiedad conmutativa de la intersección es otro axioma. Igualmente evidente es la propiedad

$$xx = x$$

axioma en el que la lógica se aparta del álgebra ordinaria. Boole aceptó también como axiomático que

$$x + y = y + x$$

y

$$x(u + v) = xu + xv.$$

Con estos axiomas, la ley del tercio excluso puede formularse ya en la forma

$$x + (1 - x) = 1$$

es decir, todo objeto es x o no x. Todo X es Y se expresa por $x(1 - y) = 0$. Ningún X es Y por $xy = 0$. Algunos X son Y por $xy \neq 0$, y algunos X no son Y por $x(1 - y) \neq 0$.

Boole planeaba deducir de los axiomas las leyes generales del razonamiento, aplicando repetidamente los procesos permitidos por dichos axiomas. Como conclusiones triviales obtuvo que $1x = x$ y que $0x = 0$. Un ejemplo de un argumento un poco más complicado es el siguiente: de

$$x + (1 - x) = 1$$

se sigue que

$$z(x + (1 - x)) = z \cdot 1$$

y por tanto que

$$zx + z(1 - x) = z.$$

Así pues, la clase de objetos z consiste en los que están en x y en los que están en $1 - x$.

Boole observó que el cálculo de clases se podía interpretar también como un cálculo de proposiciones. Así, si x e y son proposiciones en vez de clases, entonces xy sería la afirmación conjunta de x e y, y $x + y$ la afirmación de x o y o ambas. La proposición $x = 1$ significaría que x es verdadera, mientras que $x = 0$ que x es falsa; $1 - x$ significaría la negación de x. Sin embargo, Boole no llegó demasiado lejos con su cálculo de proposiciones.

Tanto a de Morgan como a Boole los podemos considerar en justicia como los reformadores de la lógica aristotélica y los iniciadores del álgebra de la lógica. El resultado de sus obras fue el de construir una ciencia de la lógica que iba a estar en adelante separada de la filosofía y unida a la matemática.

El cálculo de proposiciones progresó de la mano de Charles S. Peirce. Peirce distinguió entre una proposición y una función proposicional. Una proposición, «Juan es un hombre», sólo contiene constantes, mientras que una función proposicional, «x es un hombre», contiene variables. Mientras que una proposición es verdadera o falsa, una función proposicional en general es verdadera para algunos valores de la variable y falsa para otros. Peirce introdujo también las funciones proposicionales de dos variables, por ejemplo, «x conoce a y».

Todos los que habían contribuido a construir la lógica simbólica hasta este momento, estaban interesados principalmente en la lógica y en matematizarla. Con la obra de Gottlob Frege (1848-1925), profesor de matemáticas de Jena, la lógica matemática toma una dirección nueva, que es la que nos interesa en relación con los problemas de fundamentación de la matemática. Frege escribió varias obras importantes: *Begriffsschrift* (Cálculo de conceptos, 1879), *Die Grundlagen der Arithmetik* (Los Fundamentos de la Aritmética, 1884), y *Grundgesetze der Arithmetik* (Las Leyes Fundamentales

de la Aritmética; vol. 1, 1893; vol. 2, 1903). Todas estas obras se caracterizan por su alto nivel de precisión y completitud en los detalles.

En el campo de la lógica propiamente dicha, Frege generalizó el uso de las variables, cuantificadores y funciones proposicionales; la mayor parte de esta obra fue llevada a cabo independientemente de sus predecesores, incluido Pierce. En el *Begriffsschrift* Frege da una fundamentación axiomática de la lógica, introduciendo muchas distinciones que más tarde adquirieron gran importancia; por ejemplo, la distinción entre el enunciado de una proposición y la afirmación de que es verdadera. La afirmación de su veracidad se representa colocando el símbolo $\vdash$ delante de la proposición en cuestión. También distinguió entre un objeto x y el conjunto $\{x\}$ que contiene al elemento x solamente, y entre la pertenencia a un conjunto y la inclusión de un conjunto en otro. Utilizó variables y funciones proposicionales, lo mismo que Pierce, indicando la cuantificación de sus funciones proposicionales, es decir, el dominio de la variable o variables para la que son verdaderas. También introdujo (1879) el concepto de implicación material: A implica B significa que o bien A es verdadera y B verdadera, o A es falsa y B verdadera, o A es falsa y B falsa. Esta interpretación de la implicación es la conveniente para la lógica matemática. Frege se ocupó también de la lógica de relaciones; así, por ejemplo, la relación de orden «a es mayor que b» es muy importante en su obra.

Una vez edificada la lógica sobre un sistema explícito de axiomas, procede Frege en su *Grundlagen* a atacar el verdadero problema, el de *construir la matemática como extensión de la lógica*, expresando los conceptos de la aritmética en términos de conceptos lógicos. Así pues las definiciones y leyes de la aritmética se derivarían de premisas puramente lógicas. Examinaremos esta construcción al hablar de la obra de Russell y Whitehead. Desgraciadamente el simbolismo de Frege era muy complicado y extraño para los matemáticos, con la consecuencia de que su obra fue poco conocida, de hecho, hasta que la descubrió Russell. Por ello resulta bastante irónico el hecho de que, justo en el momento en que el segundo volumen del *Grundgesetze* estaba en prensa, recibiera Frege una carta de Russell en la que le informaba de las paradojas de la teoría de conjuntos. Al final del volumen 2 (p. 253) escribe Frege: «Difícilmente puede ocurrirle a un científico algo menos deseable que ver tambalearse los fundamentos de su obra recién terminada. Me he visto en esta posición

por una carta de Mr. Bertrand Russell cuando esta obra estaba casi terminada de imprimir.»

5. La escuela logicista

Dejamos la exposición de las investigaciones sobre los fundamentos de la matemática en el punto en el que la axiomatización de la teoría de conjuntos había suministrado una fundamentación que evitaba las paradojas conocidas y sin embargo servía de base lógica para toda la matemática existente. También señalamos que este planteamiento no resultaba satisfactorio a muchos matemáticos. Todos reconocían que quedaba por demostrar la consistencia del sistema de los números reales y de la teoría de conjuntos; tal consistencia ya no podía seguir siendo una cuestión menor y, por ejemplo, el uso del axioma de elección era muy controvertido. Pero, aparte de estos problemas, estaba la cuestión más general de cuál había de ser la fundamentación correcta de la matemática. El movimiento axiomático de finales del siglo XIX y la axiomatización de la teoría de conjuntos habían procedido sobre la base de que la lógica empleada por la matemática podía considerarse como algo dado y seguro. Pero a comienzos del siglo XX aparecieron varias escuelas de pensamiento que ya no estaban de acuerdo con esta hipótesis. La escuela encabezada por Frege intentó reconstruir la lógica y construir la matemática dentro de ella. Este plan, como hemos visto, se vio frenado por la aparición de las paradojas, pero no fue abandonado. De hecho, fue adoptado y desarrollado por Bertrand Russell y Alfred North Whitehead. Por su parte, Hilbert, consciente ya de la necesidad de establecer la consistencia, comenzó a formular sus propias ideas sistemáticas sobre la fundamentación de la matemática. Hubo aún otro grupo de matemáticos, conocidos como los intuicionistas, descontentos con los conceptos y demostraciones que se habían introducido en el análisis de finales del siglo XIX. Adoptaban una posición filosófica que no sólo no se podía reconciliar con parte de la metodología del análisis, sino que cuestionaba el papel de la lógica. El desarrollo de estas varias filosofías constituyó la tarea más importante en la fundamentación de la matemática y su resultado fue el de volver a plantear de manera completa la cuestión siempre abierta de la naturaleza de la matemática. Analizaremos por separado cada una de estas tres escuelas de pensamiento principales.

La primera de ellas se conoce como escuela logicista y su filosofía recibe el nombre de logicismo, siendo sus fundadores Russell y Whitehead. Independientemente de Frege, concibieron la idea de que la matemática es derivable exclusivamente de la lógica y, en consecuencia, es una extensión de la lógica. Las ideas básicas aparecen bosquejadas por Russell en sus *Principles of Matehematics* (1903), y se verían desarrolladas en detalle en la obra de Whitehead y Russell *Principia Mathematica* (3 vols., 1910-1913). Dado que los *Principia* constituyen la versión definitiva, nos basaremos en ella para nuestra exposición.

Esta escuela parte del desarrollo de la lógica misma, de la que se seguirá la matemática, sin necesidad de ningún axioma específicamente matemático. El desarrollo de la lógica consistirá en establecer para ella un sistema de axiomas, del que se irán deduciendo teoremas que pueden utilizarse en los razonamientos sucesivos. Así pues, las leyes de la lógica se derivarán formalmente de los axiomas. Los *Principia* parten también de conceptos indefinidos, lo mismo que cualquier teoría axiomática, puesto que no es posible definir todos los términos sin recurrir a un regreso al infinito en las definiciones. Algunos de estos conceptos indefinidos son los de proposición elemental, de función proposicional, la afirmación de la verdad de una proposición elemental, la negación de una proposición y la disyunción de dos proposiciones.

Russell y Whitehead explican estos conceptos, aunque, como indican expresamente, tales explicaciones no forman parte del desarrollo lógico. Por una proposición entienden simplemente una sentencia cualquiera que afirma un estado de cosas o una relación: por ejemplo, Juan es un hombre, las manzanas son rojas, etc. Una función proposicional es una expresión que contiene al menos una variable, de manera que al sustituir un valor por dicha variable se obtiene una proposición. Así, por ejemplo, «x es un entero» es una función proposicional. La negación de una proposición significa «No es cierto que la proposición se verifique», de manera que si p es la proposición «Juan es un hombre», entonces la negación de p, representada por $\neg p$, significa «No es cierto que Juan sea un hombre», o «Juan no es un hombre». La disyunción de dos proposiciones p y q, representada por $p \vee q$, significa que «Se verifica p o se verifica q». El significado de «*o*» aquí es el que se sobreentiende en la sentencia «Pueden solicitarlo hombres o mujeres»; es decir, los hombres pueden solicitarlo, las mujeres pueden solicitarlo, y unos y otros pueden

solicitarlo. En cambio en la sentencia «Esta persona es un hombre o una mujer», la partícula «*o*» tiene el significado más común de o bien lo uno o lo otro, pero no las dos cosas a la vez. En matemáticas se utiliza el «*o*» en el primer sentido, aunque a veces el segundo es el único posible. Por ejemplo, «El triángulo es isósceles o el cuadrilátero es un paralelogramo», ilustra el primer sentido, pero también podemos decir «Todo número no nulo es positivo o negativo», sabiendo por la aritmética de los números reales que un número no puede ser a la vez positivo y negativo. Así pues, la afirmación $p \vee q$ significa p y q o bien p y $\neg q$ o bien $\neg p$ y q.

Una relación de la mayor importancia entre proposiciones es la de implicación, es decir, que la verdad de la primera impone la verdad de la segunda. En los *Principia* la implicación $p \Rightarrow q$ viene definida por $\neg p \vee q$, lo que, a su vez, significa que se verifican $\neg p$ y q o bien $\neg p$ y $\neg q$ o bien p y q. Como ejemplo considérese la implicación «Si X es un hombre, entonces X es mortal»; aquí podrían ocurrir tres cosas:

X no es hombre y X es mortal.
X es un hombre y X es mortal.
X no es hombre y X no es mortal.

Cualquiera de estas posibilidades es admisible; lo único que la implicación prohíbe es que

X es un hombre y X no es mortal.

Algunos de los postulados de los *Principia* son:

a) Cualquier cosa implicada por una proposición elemental verdadera es verdadera.
b) $(p \vee p) \Rightarrow p$
c) $q \Rightarrow (p \vee q)$
d) $(p \vee q) \Rightarrow (q \vee p)$
e) $(p \vee (q \vee r)) \Rightarrow (q \vee (p \vee r))$.
f) La afirmación de p y la afirmación de $p \Rightarrow q$ permite la afirmación de q.

La independencia y la consistencia de estos postulados no puede demostrarse por los métodos usuales, inaplicables aquí. De estos postulados los autores proceden a deducir teoremas de la lógica y por último la aritmética y el análisis. Las reglas conocidas de los silogismos aristotélicos aparecen también como teoremas, por supuesto.

Para ilustrar cómo incluso la lógica misma ha sido formalizada deductivamente, fijémonos en unos pocos teoremas de la primera parte de los *Principia Mathematica:*

$$(p \Rightarrow \neg p) \Rightarrow \neg p \qquad 2.01$$

que es el principio de *reductio ad absurdum;* en palabras, si la hipótesis de p implica que p es falsa, entonces p es falsa.

$$(q \Rightarrow r) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) \qquad 2.05$$

que es una forma de silogismo; en palabras, si q implica r, entonces si p implica q, p implica r.

$$p \vee \neg p \qquad 2.11$$

que es el principio de tercio excluso: p es verdadera o falsa.

$$p \Rightarrow \neg(\neg p) \qquad 2.12$$

en palabras, p implica que $\neg p$ es falsa.

$$(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p) \qquad 2.16$$

Si p implica q, no-q implica no-p.

Las proposiciones constituyen una etapa hacia las funciones proposicionales, que permiten tratar conjuntos por medio de las correspondientes propiedades, en vez de nombrar sus elementos. Así, la función proposicional «x es rojo» representa el conjunto de todos los objetos rojos.

Si los elementos de un conjunto son objetos individuales, entonces la función proposicional que se les aplica se llama de tipo 0. Si los elementos de un conjunto son ellos mismos funciones proposicionales de tipo 0, cualquier función proposicional que se les aplique se llamará de tipo 1 y, en general, cualquier función proposicional cuyas variables sean de tipos menores o iguales que n, se llamará de tipo $n + 1$.

La teoría de los tipos trata de evitar las paradojas que surgen al tratar con colecciones de objetos que contienen un elemento definido solamente en términos de la colección completa. La solución

propuesta por Russell y Whitehead a esta dificultad consistía en exigir que «aquello que supone previamente todos los elementos de una colección, no puede ser un elemento de dicha colección». Para llevar a cabo esta restricción en los *Principia,* imponen que una función (lógica) no puede tener como uno de sus argumentos nada que esté definido en términos de la función misma. A continuación pasan a discutir las paradojas y demuestran que la teoría de tipos las evita.

Sin embargo, la teoría de los tipos nos conduce a clases de afirmaciones que hay que distinguir cuidadosamente según su tipo. Si se intenta construir la matemática dentro de la teoría de los tipos, el desarrollo se hace excesivamente complicado. Por ejemplo, en los *Principia* dos objetos a y b son iguales si para toda propiedad $P(x)$, $P(a)$ y $P(b)$ son proposiciones equivalentes (cada una implica la otra). Según la teoría de tipos, P puede ser de diferentes tipos, porque puede contener variables de varios niveles además de los objetos a o b, y así la definición de igualdad debe aplicarse a todos los posibles tipos de P; en otras palabras, hay infinitas relaciones de igualdad, una para cada tipo de la correspondiente propiedad. Análogamente, un número irracional definido por una cortadura de Dedekind es un objeto de un tipo más alto que un número racional, que a su vez es de un tipo más alto que un número natural, y así el continuo consiste de números de diferentes tipos. Para escapar a esta complejidad introdujeron Russell y Whitehead el llamado «axioma de reducibilidad», que afirma la existencia, para cada función proposicional de cualquier tipo, de una función proposicional equivalente de tipo cero.

Una vez estudiadas las funciones proposicionales, los autores se ocupan de la teoría de clases. Una clase, dicho un poco vagamente, es la colección de objetos que satisfacen alguna función proposicional. Las relaciones se consideran entonces como clases de parejas que satisfacen funciones proposicionales de dos variables; así, «x juzga a y» expresa una relación. Sobre esta base los autores están preparados ya para introducir el concepto de número cardinal.

La definición de número cardinal tiene un gran interés, dependiendo de la relación de correspondencia biunívoca entre clases introducida previamente. Si dos clases están en correspondencia biunívoca se dice que son semejantes; la relación de semejanza es reflexiva, simétrica y transitiva. Todas las clases semejantes tienen una propiedad común, que es su número de elementos. Sin embargo, las clases semejantes bien podrían tener más de una propiedad común. Para evitar esta dificultad Russell y Whitehead definieron el número

de elementos de una clase, como ya había hecho Frege, como la clase de todas las clases semejantes a la clase dada. Así, el número tres es la clase de todas las clases con tres elementos, es decir, de todas las clases x, y, z tales que $x \neq y \neq z$. Dado que la definición de número presupone el concepto de correspondencia biunívoca, podría parecer que tal definición es circular. Los autores hacen notar, sin embargo, que una relación es biunívoca cuando, si x y x' están relacionados con y, entonces x y x' han de ser idénticos, y viceversa, si x está relacionado con y e y', entonces y e y' han de ser idénticos. Por tanto, es claro que el concepto de correspondencia biunívoca en realidad no hace uso del número 1.

Dados los números cardinales y los naturales, es posible construir ya los sistemas de los números reales y complejos, las funciones y, de hecho, todo el análisis. La geometría puede introducirse a través de los números. Aunque los detalles en los *Principia* difieren algo de nuestra propia exposición de los fundamentos de los distintos sistemas numéricos y de la geometría (caps. 41 y 42), es fácil ver que tales construcciones son lógicamente posibles sin axiomas adicionales.

Este era, pues, el grandioso programa de la escuela logicista. Lo que hizo por la lógica misma fue notable, como acabamos de esbozar brevemente. Lo que hizo por la matemática, y esto hemos de subrayarlo claramente, fue fundamentar la matemática en la lógica. No era necesario ningún axioma propio de la matemática; la matemática se reducía así a una extensión natural de las leyes y contenidos de la lógica. Pero los postulados de la lógica y todas sus consecuencias tienen un carácter arbitrario y además formal. Es decir, no tienen contenido propio, sólo tienen forma. Como consecuencia, la matemática tampoco tiene contenido, sino sólo forma. El significado físico que atribuimos a los números o a los conceptos geométricos no forman parte de la matemática. Esto es lo que tenía Russell en la mente cuando dijo que la matemática es la materia en la que nunca sabemos de qué estamos hablando ni si lo que decimos es verdad. En realidad, cuando Russell comenzó a desarrollar su programa a principios de siglo, pensaba (como Frege) que los axiomas de la lógica eran verdades, pero abandonó este punto de vista desde la edición de 1937 de los *Principles of Mathematics*.

El planteamiento logicista ha sido muy criticado. Concretamente, el axioma de reducibilidad levantó considerable oposición, por su carácter bastante arbitrario. Ha sido llamado un feliz accidente y no

una necesidad lógica; se ha dicho que tal axioma no tiene lugar en la matemática, y que lo que no puede demostrarse sin él, no puede considerarse como demostrado en absoluto; otros llaman a este axioma un verdadero sacrificio del intelecto. Además, el sistema de Russell y Whitehead nunca se llegó a completar, y es oscuro en numerosos detalles. Más tarde se hicieron muchos esfuerzos por simplificarlo y clarificarlo.

Otra crítica filosófica seria de la posición logicista en su totalidad es la de que, si el punto de vista logicista es correcto, entonces toda la matemática es una ciencia lógico-deductiva puramente formal, cuyos teoremas se siguen de las leyes del pensamiento exclusivamente. Entonces parece necesario explicar cómo es posible que tal elaboración deductiva de las leyes del pensamiento sirva para representar tan gran diversidad de fenómenos naturales como la acústica, el electromagnetismo, la mecánica, etc. Por otra parte, en la creación de la matemática una intuición sensorial o imaginativa debe suministrar nuevos conceptos, sean derivados de la experiencia o no. De otra manera, ¿cómo podrían surgir conocimientos nuevos? Pero en los *Principia* todos los conceptos se reducen a conceptos lógicos.

La formalización del programa logicista no parece representar la matemática en ningún sentido real. Nos presenta la cáscara pero no la sustanciosa semilla. Poincaré llegó a decir, maliciosamente (*Les Fondements de la Science*, p. 483), «La teoría logicista no es estéril; engendra contradicciones». Esto no es cierto si uno acepta la teoría de tipos, pero esta teoría, como hemos dicho, resulta artificial. Weyl también atacó al logicismo, diciendo que su compleja estructura «pone a prueba la fuerza de nuestra fe apenas menos que las doctrinas de los primeros Padres de la Iglesia o de los filósofos escolásticos de la Edad Media».

A pesar de esta crítica, muchos matemáticos siguen aceptando la filosofía logicista. La construcción llevada a cabo por Russell y Whitehead tuvo importantes consecuencias en otra dirección, al conducir a una axiomatización completa de la lógica en forma totalmente simbólica, lo cual permitió a la lógica matemática hacer enormes progresos.

6. La escuela intuicionista

Otro grupo de matemáticos, más tarde llamados intuicionistas, se plantearon un enfoque radicalmente distinto de la matemática. Al igual que en el caso del logicismo, la filosofía intuicionista tuvo sus orígenes a finales del siglo XIX, cuando la rigorización del sistema numérico y de la geometría eran tareas de la mayor importancia. El posterior descubrimiento de las paradojas de la teoría de conjuntos vino a impulsar su desarrollo.

El primer intuicionista fue Kronecker, que expresó sus puntos de vista durante las décadas de los 1870 y 1880. Para Kronecker, el rigor impuesto por Weierstrass involucraba conceptos inaceptables, y la obra de Cantor sobre teoría de conjuntos y números transfinitos no era matemática sino misticismo. Kronecker estaba dispuesto a aceptar los números enteros porque son claros a la intuición. Estos eran «obra de Dios»; todo lo demás era obra del hombre y, por tanto, sospechoso. En su ensayo de 1887, «Uber den Zahlbegriff» (Sobre el concepto de número),[12] mostró cómo algunos tipos de números, por ejemplo los números racionales, se podrían definir en términos de números enteros. Los números fraccionarios como tales eran aceptables por un convenio de notación. Kronecker quería prescindir de la teoría de números irracionales y funciones continuas; su ideal era que todo teorema del análisis se pudiera interpretar en términos de relaciones entre enteros únicamente.

Otra objeción que presentaba Kronecker a muchas partes de la matemática era la de que no daban métodos o criterios constructivos para determinar en un número finito de pasos los objetos que manejaban. Las definiciones deberían incluir los medios necesarios para calcular el objeto definido en un número finito de pasos, y las demostraciones de existencia deberían permitir el cálculo del número cuya existencia se demostraba con un grado de aproximación arbitrario. Los algebristas solían contentarse con decir de un polinomio $f(x)$ que tenía una raíz racional, en cuyo caso $f(x)$ era reducible, y en caso contrario irreducible. En su *Festschrift* «Grundzüge einer arithmetischen Theorie der algebraischen Grössen» (Elementos de una teoría aritmética de magnitudes algebraicas) [13] decía Kronecker, «La definición de reducibilidad está desprovista de fundamento se-

[12] Jour. für Math., 101, 1887, 337-355 = *Werke*, 3, 251-274.
[13] Jour. für Math., 92, 1882, 1-122 = *Werke*, 2, 237-387.

guro mientras no se dé un *método* en virtud del cual se pueda decidir si una función dada es reducible o no».

En otro contexto, a pesar de que las diversas teorías de números irracionales definían cuándo dos números reales a y b son iguales, o $a > b$ o $b > a$, no daban criterios para determinar qué alternativa se verificaba en un caso dado. Por tanto Kronecker planteó objeciones a tales definiciones; para él sólo eran definiciones en apariencia. La teoría de números irracionales le parecía insatisfactoria en su totalidad, y así un día le dijo a Lindemann, que acababa de demostrar que π era un irracional trascendente: «¿De qué sirve su bella investigación sobre π? ¿Para qué estudiar tales problemas, si los números irracionales no existen?».

Kronecker mismo no hizo mucho por desarrollar la teoría intuicionista, excepto criticar la ausencia de métodos constructivos para determinar objetos matemáticos cuya existencia simplemente se había demostrado. Intentó reconstruir el álgebra pero no hizo ningún esfuerzo por reconstruir el análisis. Por otra parte, Kronecker realizó una importante obra en aritmética y álgebra que no se ajusta a sus propias exigencias porque, como observó Poincaré,[14] a veces olvidaba temporalmente su propia filosofía.

En su época Kronecker no encontró partidarios de su filosofía, y durante casi veinticinco años nadie siguió sus ideas. Sin embargo, una vez descubiertas las paradojas, el intuicionismo se revitalizó y se convirtió en un movimiento serio y muy extendido. El siguiente y enérgico defensor fue Poincaré. Ya hemos mencionado su oposición a la teoría de conjuntos porque conducía a las paradojas, y tampoco aceptó el programa logicista para salvar la matemática. Ridiculizó los intentos de basar la matemática en la lógica porque reduciría la matemática a una inmensa tautología. También se burló de la muy artificial (para él) introducción de los números; así, en los *Principia* se define el 1 como $\hat{\alpha}\{\exists x \cdot \alpha = \iota' x\}$, ante lo cual Poincaré decía sarcásticamente que esta era una definición admirable para dársela a gente que nunca hubiera oído hablar del número 1.

En su obra *Science et Méthode (Les Fondaments de la Science,* p. 480) dice:

> El logicismo tiene que ser reconstruido, y uno no puede estar demasiado seguro de lo que se pueda salvar. Es innecesario añadir que sólo se ponen

---

[14] Acta Math., 22, 1899, 17.

en cuestión el cantorismo y el logicismo; la verdadera matemática, la que sirve a algún fin útil, puede seguir desarrollándose de acuerdo con sus propios principios, sin prestar atención alguna a las tempestades desencadenadas en el exterior, y continuará sus conquistas usuales, etapa tras etapa, conquistas que son definitivas y que no necesitará abandonar nunca.

Poincaré rechazaba los conceptos que no se pueden definir con un número finito de palabras; por ejemplo, un conjunto construido con ayuda del axioma de elección, no está realmente definido cuando se ha elegido un elemento de cada uno de un número infinito de conjuntos. También sostenía que la aritmética no puede justificarse por ninguna fundamentación axiomática. Nuestra intuición es anterior a tal estructura y, en particular, la inducción completa o matemática procede de una intuición fundamental y no se reduce a un axioma que casualmente es útil en algunos sistemas axiomáticos. Al igual que Kronecker, Poincaré insistía en que todas las definiciones y demostraciones tenían que ser constructivas.

Poincaré estaba de acuerdo con Russell en que el origen de las paradojas estaba en la definición de colecciones o conjuntos que incluían el objeto definido. Así, el conjunto A de todos los conjuntos contiene a A como elemento, pero A no puede definirse mientras no esté definido cada elemento de A, y si A es uno de ellos la definición es circular. Otro ejemplo de definición impredicativa es la del valor máximo de una función continua sobre un intervalo cerrado. Tales definiciones eran frecuentes en análisis y especialmente en la teoría de conjuntos.

Otras críticas a la situación lógica de la matemática de la época se desarrollaron y discutieron en un intercambio epistolar entre Borel, Baire, Hadamard y Lebesgue.[15] Borel defendía la afirmación de Poincaré de que los números naturales no pueden fundamentarse axiomáticamente, y criticaba además al axioma de elección porque exige una infinitud no numerable de elecciones simultáneas, lo cual es inconcebible para la intuición. Hadamard y Lebesgue iban más lejos, sosteniendo que incluso una infinitud numerable de elecciones arbitrarias sucesivas no es más intuitiva porque sigue exigiendo una infinitud de operaciones, que son imposibles de concebir como realizadas de manera efectiva. Para Lebesgue todas las dificultades se reducían a saber lo que uno entiende al decir que un objeto mate-

---

[15] Ver nota 9.

mático existe. En el caso del axioma de elección sostenía que si uno simplemente «piensa» en una manera de elegir, ¿no puede entonces cambiar sus elecciones en el curso del razonamiento? Incluso la elección de un único elemento de un conjunto no vacío planteaba, según Lebesgue, las mismas dificultades; uno tiene que saber que el objeto «existe», lo que significa que debe indicar la elección explícitamente. Así pues, Lebesgue rechazaba la demostración de Cantor de la existencia de números transcedentes. Hadamard indicó que las objeciones de Lebesgue conducían a negar la existencia del conjunto de los números reales, y Borel llegó exactamente a la misma conclusión.

Todas las objeciones de los intuicionistas que hemos mencionado fueron esporádicas y fragmentarias. El fundador sistemático del intuicionismo moderno es Brouwer. Al igual que Kronecker, la mayor parte de su obra matemática, especialmente en topología, no estaba de acuerdo con su filosofía, pero ello no implica la menor duda sobre la seriedad de su posición. Desde su tesis doctoral, *Sobre los Fundamentos de la Matemática* (1907), comenzó Brouwer a construir su filosofía intuicionista, y desde 1918 en adelante desarrolló y generalizó sus puntos de vista en una serie de artículos publicados en diversas revistas, incluidos los *Mathematische Annalen* de 1925 y 1926.

La posición intuicionista de Brouwer se deriva de una filosofía más general. La intuición fundamental, según él, es la presencia de percepciones en una sucesión temporal. «La matemática surge cuando la cuestión de la "paridad", que resulta del paso del tiempo, se abstrae de todas las apariciones concretas. La forma vacía que permanece [la relación de n a $n + 1$] del contenido común de todas estas "paridades" se convierte en la intuición original de la matemática, y repetida indefinidamente crea nuevos objetos matemáticos». Así pues, por repetición ilimitada construye la mente el concepto de la sucesión de los números naturales. Esta idea de que los números naturales se derivan de la intuición del tiempo la habían defendido ya Kant, William R. Hamilton en su «Algebra as a Science of Time», y el filósofo Arthur Schopenhauer.

Brouwer concibe el pensamiento matemático como un proceso de construcción que edifica su propio universo independiente del de nuestra experiencia y algo así como un modelo libre, restringido únicamente en tanto que está basado en la intuición matemática fundamental. Este concepto intuitivo fundamental no debe ser concebido como una idea indefinida, tal como ocurre en las teorías axio-

máticas, sino algo más bien como algo en *términos de lo cual* han de ser concebidas intuitivamente todas las ideas indefinidas que aparecen en los diversos sistemas matemáticos, si es que han de servir realmente en el pensamiento matemático.

Brouwer sostiene que «en este proceso constructivo limitado por la obligación de reconocer con cuidado qué tesis son aceptables a la intuición y cuáles no, yace la única fundamentación posible de la matemática». Las ideas matemáticas están en la mente humana *previamente al lenguaje, la lógica y la experiencia.* Es la intuición, y no la experiencia ni la lógica la que determina la validez y aceptabilidad de las ideas. Hay que recordar, desde luego, que estas afirmaciones sobre el papel de la experiencia hay que tomarlas en el sentido filosófico y no en el histórico.

Para Brouwer los objetos matemáticos los adquirimos por construcción intelectual, donde los números básicos 1, 2, 3, ... dan el prototipo de tales construcciones. La posibilidad de la repetición ilimitada de la forma vacía, es decir, la etapa de n a $n + 1$ conduce a los conjuntos infinitos. Sin embargo, el infinito de Brouwer es el infinito potencial de Aristóteles, mientras que la matemática moderna tal como la fundamenta Cantor, por ejemplo, hace un amplio uso de conjuntos infinitos en sentido actual, cuyos elementos están presentes «todos de golpe».

Con respecto a la idea intuicionista de conjunto infinito, dice Weyl, que perteneció a la escuela intuicionista, que

> ... la sucesión de los números que crece más allá de cualquier nivel ya alcanzado... es una variedad de posibilidades que se abre al infinito; permanece para siempre en el status de creación, pero no es un dominio cerrado de cosas que existan por sí mismas. El haber convertido ciegamente lo uno en lo otro constituye el verdadero origen de nuestras dificultades, incluyendo las antinomias, un origen de naturaleza más fundamental que el principio del círculo vicioso de Russell, ya indicado. Brouwer nos abrió los ojos y nos mostró hasta dónde la matemática clásica, alimentada por una creencia en lo absoluto que trasciende todas las posibilidades de realización humanas, se remonta más allá de las afirmaciones que pueden pretender tener un significado real y una verdad fundada en la evidencia.

El mundo de la intuición matemática se opone al mundo de las percepciones causales. A este mundo causal, y no a la matemática, corresponde el lenguaje, que sirve allí para el entendimiento en los asuntos comunes; las palabras o expresiones verbales se utilizan para

comunicar verdades. El lenguaje sirve para evocar copias de ideas en las mentes de los hombres por medio de símbolos y sonidos. Pero los pensamientos nunca pueden ser simbolizados completamente, y estas observaciones se aplican igualmente al lenguaje matemático, incluido el lenguaje simbólico. Las ideas matemáticas son independientes de la vestidura del lenguaje y, de hecho, mucho más ricas.

La lógica se refiere al lenguaje, presentando un sistema de reglas que permiten deducir unas conexiones verbales de otras y que también intentan comunicar verdades. Sin embargo, estas últimas verdades no son tales antes de que se tenga experiencia de ellas, y tampoco está garantizado que se pueda alcanzar dicha experiencia. La lógica no es un instrumento seguro para descubrir verdades, y no puede deducir verdades que no se puedan obtener de alguna otra manera. Los principios lógicos expresan la regularidad observada a posteriori en el lenguaje, y se reducen a un instrumento para manipular el lenguaje o una teoría de la representación de dicho lenguaje. Los progresos más importantes de la matemática no se obtienen perfeccionando la forma lógica sino modificando la teoría básica misma. La lógica se apoya en la matemática y no la matemática en la lógica.

Dado que Brouwer no reconoce ningún principio lógico obligatorio a priori, tampoco reconoce la tarea matemática de deducir conclusiones de axiomas. La matemática no está obligada a respetar las reglas de la lógica, y por este motivo las paradojas carecen de importancia incluso si tuviéramos que aceptar los conceptos y construcciones matemáticas en que aparecen. Desde luego, como veremos, los intuicionistas no aceptan todos estos conceptos y demostraciones.

Sobre el papel de la lógica dice Weyl:[16]

Según su punto de vista [el de Brouwer] y lectura de la historia, la lógica clásica resultó por abstracción de la matemática de conjuntos finitos y sus subconjuntos... Olvidando este origen limitado se aplicó equivocadamente, más tarde, esa misma lógica para algo anterior a toda la matemática y por encima de ella, y por último se terminó aplicando, sin justificación, a la matemática de los conjuntos infinitos. En esto consiste la caída y el pecado original de la teoría de conjuntos, por el cual resulta justamente castigada

[16] Amer. Math. Monthly, 53, 1946, 2-13 = *Ges. Abh.*, 4, 268-279.

en las antinomias. Lo sorprendente no es que aparecieran tales contradicciones, sino que lo hicieran en una etapa tan avanzada del juego.

En el campo de la lógica hay algunos principios o procedimientos claros, *intuitivamente* aceptables, que pueden ser utilizados para obtener nuevos teoremas de otros anteriores. Estos principios forman parte de la intuición matemática fundamental. Sin embargo, no todos los principios lógicos son aceptables por la intuición básica, y debería someterse a crítica lo que ha sido aceptado desde la época de Aristóteles. Las antinomias se han producido porque los matemáticos han aplicado libremente esas leyes aristotélicas. Por tanto los intuicionistas tenían que proceder a un análisis acerca de qué principios lógicos son legítimos para que la lógica usual esté de acuerdo con las intuiciones correctas y sea capaz de expresarlas.

Como ejemplo concreto de un principio lógico que se aplica demasiado libremente, menciona Brouwer la ley de tercio excluso. Este principio, que afirma que toda proposición significativa es verdadera o falsa, es fundamental para el método de demostración indirecta. Históricamente surgió por aplicación de razonamientos a subconjuntos de conjuntos finitos, por abstracción. Fue aceptado entonces como un principio independiente a priori y se aplicó injustificadamente a conjuntos infinitos. Mientras que en el caso de conjuntos finitos es posible saber si todos los elementos tienen una propiedad comprobándolo con cada uno, eso no puede hacerse con los infinitos. Puede ocurrir que sepamos que un elemento del conjunto infinito no posee una propiedad, o puede que por la misma construcción del conjunto sepamos o podamos demostrar que todo elemento tiene esa propiedad. En cualquier caso, no se puede usar la ley de tercio excluso para demostrar que la propiedad se verifica.

Por tanto, si uno demuestra que no todos los elementos de un conjunto infinito tienen una propiedad, entonces Brouwer rechaza que se haya demostrado la conclusión de que existe por lo menos un elemento que no tiene dicha propiedad. Así, por ejemplo, de la negación de que se verifique $a^b = b^a$ para todos los números, los intuicionistas no concluyen que existan a y b tales que $a^b \neq b^a$. Como consecuencia, muchas demostraciones clásicas de existencia no son aceptadas por los intuicionistas. La ley de tercio excluso puede ser utilizada en los casos en que la conclusión pueda alcanzarse en un número finito de etapas; por ejemplo, para decidir la cuestión de si

un libro contiene erratas. En cualquier otro caso los intuicionistas rechazan la posibilidad de una decisión.

El rechazo de la ley de tercio excluso da lugar a una nueva posibilidad, la de proposiciones indecidibles. Los intuicionistas sostienen, con respecto a los conjuntos *infinitos*, que hay una tercera posibilidad, a saber, que una proposición no sea ni demostrable ni refutable. Como ejemplo de tal proposición, definamos k como el lugar que ocupa el primer cero que aparece seguido por las cifras 1, 2, ..., 9 en el desarrollo decimal de π. La lógica aristotélica nos dice que k existe o no existe, y los matemáticos que siguen a Aristóteles pueden razonar a partir de estas dos posibilidades. Brouwer, en cambio, rechaza todos estos razonamientos debido a que no sabemos si podremos demostrar que k existe o no existe. Así pues, hay cuestiones matemáticas que pueden no ser decididas nunca a partir de las afirmaciones que expresan los axiomas de la matemática. Tales cuestiones pueden parecernos decidibles, pero en realidad nuestra base para esperar que esto ocurra se reduce a que tratan de conceptos matemáticos.

Con respecto a los conceptos aceptables como legítimos para una discusión matemática, los intuicionistas insisten en que han de tener definiciones constructivas. Para Brouwer, como para todos los intuicionistas, el infinito existe exactamente en el sentido de que uno puede encontrar siempre un conjunto finito mayor que otro dado. Para discutir cualquier otro tipo de infinito, los intuicionistas exigirían que se dé un método de construir o definir tal infinito en un número finito de pasos. Así, Brouwer rechaza los conjuntos infinitos de la teoría de conjuntos.

La exigencia de constructibilidad es otra que excluye cualquier concepto cuya existencia se haya establecido por un razonamiento indirecto, es decir, argumentando que la no existencia conduce a una contradicción. Aparte del hecho de que la demostración de existencia pueda utilizar la rechazable ley de tercio excluso, tal demostración no es satisfactoria para el intuicionista porque exige una definición constructiva del objeto cuya existencia se está demostrando. Tal definición constructiva debe permitir determinar el objeto con cualquier grado de aproximación deseado en un número finito de pasos. La demostración de Euclides de la existencia de infinitos números primos (cap. 4, sec. 7) no es constructiva, no permite determinar el n-ésimo primo, y por tanto no es aceptable. Si uno demostrase simplemente la existencia de enteros x, y, z, n tales que x^n +

$y^n = z^n$, ningún intuicionista aceptaría la demostración. Por otra parte, la definición de número primo sí es constructiva, pues se puede aplicar para determinar si un número lo es o no en un número finito de pasos. La insistencia en las definiciones constructivas se aplica especialmente a los conjuntos infinitos; por ejemplo, un conjunto construido aplicando el axioma de elección a infinitos conjuntos, no sería aceptable.

Weyl dice de las demostraciones de existencia no constructivas (*Philosophy of Mathematics and Natural Science*, p. 51) que nos informan de que en mundo hay un tesoro sin descubrirnos su localización. La demostración a partir de postulados no puede reemplazar a la construcción sin pérdida de significado y de valor. También señala que adherirse a la filosofía intuicionista significa abandonar muchos teoremas de existencia del análisis clásico, por ejemplo, el teorema de Weierstrass-Bolzano: un conjunto infinito acotado de números reales no tiene necesariamente un punto límite. Para los intuicionistas, si una función de una variable real existe en su sentido, entonces es *ipso facto* continua. La inducción transfinita y sus aplicaciones al análisis y la mayor parte de la teoría de Cantor se ven condenadas categóricamente. El análisis, dice Weyl, está construido sobre arena.

Brouwer y su escuela no se han limitado a ejercer la crítica, sino que han tratado de construir una nueva matemática basada en las construcciones que aceptan. Han tenido éxito salvando el cálculo infinitesimal con sus procesos de límite, pero sus construcciones son muy complicadas; también han reconstruido partes elementales del álgebra y la geometría. Al revés que Kronecker, Weyl y Brouwer permiten algunos tipos de números irracionales. Es evidente que la matemática de los intuicionistas tiene que diferir radicalmente de lo que los matemáticos habían aceptado de una manera casi universal antes de 1900.

7. La escuela formalista

La tercera de las principales filosofías de la matemática recibe el nombre de escuela formalista y su creador fue Hilbert, que comenzó a trabajar en esta línea filosófica en 1904. Sus motivaciones en esa época eran las de establecer una base para el sistema numérico sin utilizar la teoría de conjuntos, y a continuación demostrar la con-

sistencia de la aritmética. La consistencia de esta última teoría era un problema abierto de importancia vital, puesto que él mismo había demostrado que la consistencia de la geometría se reducía a la de la aritmética. También trataba de combatir la opinión de Kronecker de que era necesario excluir los números irracionales de la matemática. Hilbert aceptaba sin reservar el infinito actual y elogiaba la obra de Cantor (cap. 41, sec. 9), deseando conservar el infinito, las demostraciones puramente existenciales y los conceptos tales como el de extremo superior, cuya definición parecía ser circular.

Hilbert presentó un artículo sobre sus puntos de vista en el Congreso Internacional de Matemáticos de 1904.[17] Durante los siguientes quince años no volvió sobre el tema pero, al fin, movido por el deseo de responder a las críticas intuicionistas del análisis clásico, volvió a los problemas de fundamentos y continuó trabajando en ellos durante el resto de su carrera científica. Durante la década de los años veinte publicó varios artículos clave, y un cierto número de matemáticos fueron adoptando gradualmente sus ideas.

Su filosofía de la etapa madura incluye varias doctrinas. Siguiendo la nueva tendencia de que cualquier fundamentación de la matemática debe tener en cuenta el papel de la lógica, los formalistas sostienen que la lógica tiene que ser tratada simultáneamente con la matemática, que tiene diversas ramas, cada una de ellas con su fundamentación axiomática propia. Esta fundamentación incluye conceptos y principios tanto lógicos como matemáticos. La lógica es un lenguaje simbólico que expresa mediante fórmulas las proposiciones matemáticas y reduce el razonamiento a un proceso formal en el que unas fórmulas se siguen de otras de acuerdo con las reglas deductivas. Todos los símbolos se consideran desprovistos de significado. En su artículo de 1926 [18] dice Hilbert que los objetos del pensamiento matemático son los símbolos mismos; los símbolos son ahora la esencia y ya no representan objetos físicos idealizados. Las fórmulas pueden expresar afirmaciones intuitivamente significativas, pero estos significados no son parte de la matemática.

Hilbert conservó la ley de tercio excluso por encontrarla necesaria para la construcción del análisis. Decía,[19] «Prohibirle a un ma-

---

[17] *Proc. Third Internat. Congress of Math.*, Heidelberg, 1904, 174-185 = *Grundlagen der Geom.*, 7.ª ed., 247-261; traducción inglesa en el Monist, 15, 1905, 338-352.

[18] Math. Ann., 95, 1926, 161-190 = *Grundlagen der Geometrie*, 7.ª ed., 262-288. Ver nota 20.

[19] Weyl, Amer. Math. Soc. Bull., 50, 1944, 637 = *Ges. Abh.*, 4, 157.

temático usar el principio de tercio excluso es como prohibirle a un astrónomo usar su telescopio o a un boxeador el uso de sus puños». Puesto que la matemática trabaja solamente con expresiones simbólicas, a estas expresiones formales se les puede aplicar todas las reglas de la lógica aristotélica. En este nuevo sentido la matemática de conjuntos infinitos resultaría posible. Hilbert esperaba también eliminar las paradojas evitando el uso explícito de la palabra «todo».

Para formular los axiomas lógicos introduce Hilbert un simbolismo para los conceptos y relaciones tales como «*y*», «*o*», «no», «existe», etc. Afortunadamente el cálculo lógico (o lógica simbólica) se había desarrollado ya para otros fines y, como Hilbert mismo dice, tenía a mano lo que necesitaba. Todos los símbolos anteriores constituyen el material para construir las expresiones formales o fórmulas.

Para manejar el infinito, Hilbert usa, aparte de los axiomas ordinarios indiscutidos, el axioma transfinito.

$$A(\tau A) \rightarrow A(a)$$

que significa: si un predicado A se verifica para el objeto fiduciario τA, entonces se verifica para todos los objetos a. Así, si suponemos que A expresa «ser corruptible», entonces si Arístides el Justo es fiduciario y corruptible, todo el mundo es corruptible.

Una demostración matemática consistirá en el siguiente proceso: la afirmación de una fórmula; la afirmación de que esta fórmula implica otra; la afirmación de esta segunda fórmula. Una sucesión de etapas tales como esta, en la que las fórmulas afirmadas vienen precedidas por axiomas o conclusiones previas, constituye la demostración de un teorema. Otra operación permisible es la sustituir un símbolo por otro o por un grupo; así pues, las fórmulas se derivan unas de otras aplicando las reglas de manipulación de símbolos en ellas.

Una proposición es verdadera si y sólo si puede obtenerse como la última de una sucesión de proposiciones, cada una de las cuales o bien es un axioma del sistema formal o se deriva de otras anteriores por una de las reglas de deducción. Cualquiera puede comprobar si una proposición dada ha sido obtenida correctamente de una sucesión de proposiciones. Así pues, bajo el punto de vista formalista, verdad y rigor están bien definidos y son objetivos.

Para el formalista, entonces, la matemática propiamente dicha es

una colección de sistemas formales, construyendo cada uno su propia lógica a la vez que su matemática, y cada uno de ellos con sus propios conceptos, sus propios axiomas, reglas deductivas tales como la igualdad y sustitución, y sus propios teoremas. El desarrollo de cada uno de estos sistemas deductivos constituye la tarea de la matemática. La matemática se convierte pues, no en una teoría sobre algo, sino en una colección de sistemas formales, en cada uno de los cuales ciertas expresiones formales se obtienen de otras por transformaciones también formales. Esto, en lo que se refiere a la parte del programa de Hilbert que se ocupa de la matemática propiamente dicha.

Sin embargo, todavía tenemos que preguntarnos si las deducciones están libres de contradicciones. Esto no necesita observarse de modo necesario intuitivamente, pero para demostrar la no contradicción, todo lo que necesitamos demostrar es que nunca se puede llegar a una expresión formal como la $0 = 1$. (Dado que, por un teorema de lógica, cualquier otra proposición se sigue de ésta y de la $0 \neq 1$, podemos limitarnos a ésta).

Hilbert y sus discípulos Wilhem Ackermann (1896-1962), Paul Bernays (1888) y von Neumann desarrollaron, durante los años 1920 y 1930 lo que Hilbert llamó *Beweistheorie* (Teoría de la Demostración) o metamatemática, un método que pretendía establecer la consistencia de un sistema formal. Hilbert proponía utilizar en la metamatemática únicamente una lógica especial básica y libre de todo tipo de objeciones. Para ello utilizó razonamientos concretos y finitos de un tipo admitido universalmente y muy próximo a los principios intuicionistas. No deberían utilizarse principios controvertidos tales como la demostración de existencia por reducción al absurdo, la inducción transfinita ni el axioma de elección, sino que las demostraciones de existencia deberían ser constructivas. Dado que un sistema formal puede ser indefinido, la metamatemática debe considerar conceptos y cuestiones en los que aparecen sistemas infinitos, al menos potencialmente. No obstante, los métodos de demostración utilizados deben ser totalmente finitarios. No debe haber ninguna referencia a una cantidad infinita de propiedades estructurales de las fórmulas ni a un número infinito de manipulaciones con ellas.

Ahora bien, la consistencia de una gran parte de la matemática clásica se puede reducir a la de la aritmética de los números naturales (o teoría de números) tal como aparece formulada en los axiomas de Peano, o a la de una teoría de conjuntos lo suficientemente rica

como para que permita demostrar dichos axiomas. Por tanto, la consistencia de la aritmética de los números naturales pasó a ser el centro de atención.

Hilbert y su escuela consiguieron demostrar la consistencia de algunos sistemas formales sencillos y creyeron estar a punto de conseguir la meta de demostrar la consistencia de la aritmética y de la teoría de conjuntos. En su artículo «Uber Unendliche» [20] dice Hilbert,

> En geometría y en la teoría física la demostración de consistencia se consigue reduciéndola a la consistencia de la aritmética. Este método falla obviamente para la demostración de la consistencia de la aritmética misma. Dado que nuestra teoría de la demostración... hace esta última etapa posible, constituye la piedra angular necesaria en la estructura de la matemática. Y, en particular, lo que ya hemos experimentado dos veces, primero en las paradojas del cálculo infinitesimal y más tarde en la teoría de conjuntos, no puede volver a ocurrir dentro del campo de la matemática.

Pero en este momento entra en escena Kurt Gödel (1906-1978). El primer trabajo importante de Gödel fue su artículo «Uber formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme *I*».[21] En él demuestra Gödel que la consistencia de un sistema que incluya la lógica usual y la teoría de números no puede ser demostrada si se limita uno a conceptos y métodos que puedan ser representables formalmente en el sistema de la teoría de números. Lo que esto significa, en efecto, es que la consistencia de la aritmética no puede ser demostrada por la lógica más restringida admisible en la metamatemática. A propósito de este resultado decía Weyl que Dios existe porque la matemática es consistente, y el diablo existe porque no podemos demostrar dicha consistencia.

El resultado anterior de Gödel es un corolario de su más sorprendente teorema. El teorema más importante de Gödel, o teorema de incompletitud, afirma que si una teoría formal axiomatizable T que incluya la aritmética es consistente, entonces T es incompleta, es decir hay una afirmación S tal que ni S ni no-S son teoremas de

---

[20] Math. Ann., 95, 1926, 161-190 = *Grundlagen der Geometrie*, 7.ª ed., 262-288. Hay traducción inglesa en Paul Benacerraf y Hilary Putnam, *Philosophy of Mathematics*, 134-181, Prentice-Hall, 1964.

[21] Monatshefte für Mathematik und Physik, 38, 1931, 173-198, véase la bibliografía.

la teoría *T*. Ahora bien, o *S* o no-*S* es verdadera; tenemos pues una afirmación verdadera que no es demostrable en la teoría. Este resultado se aplica al sistema de Russell-Whitehead, al de Zermelo-Fraenkel y a la axiomatización de Hilbert de la teoría de números. No deja de ser irónico que Hilbert, en su comunicación al Congreso Internacional de Bolonia de 1928 (ver nota 22), hubiera criticado las viejas demostraciones de completitud basadas en la categoricidad (cap. 42, sec. 3) mientras que tenía gran confianza en la completitud de su propio sistema. En realidad las viejas demostraciones sobre sistemas que contenían los números naturales solamente se habían aceptado como válidas porque la teoría de conjuntos aún no había sido axiomatizada, y se utilizaba de manera intuitiva.

El fenómeno de la incompletitud constituye un importante defecto porque entonces el sistema formal no es adecuado para demostrar todas las afirmaciones que podrían serlo correctamente (sin contradicción) dentro del sistema. Para poner las cosas peor, resulta haber afirmaciones indecidibles pero intuitivamente verdaderas en algún modelo del sistema. La incompletitud no puede remediarse añadiendo *S* o $\neg S$ como axioma, puesto que lo que demostró Gödel es que *cualquier* sistema que incluya a la aritmética contiene irremediablemente una proposición indecidible. Así pues, mientras Brouwer sostenía que lo que es intuitivamente verdadero se queda corto con respecto a lo que es matemáticamente demostrable, Gödel viene a demostrar que lo intuitivamente verdadero va más allá de la capacidad de la demostración matemática.

Una de las consecuencias del teorema de Gödel es la de que no hay ningún sistema de axiomas adecuado, no sólo para toda la matemática, sino incluso para casi cualquier rama importante de la matemática, porque cualquier tal sistema axiomático sería incompleto. Existen pues afirmaciones sobre los conceptos del sistema, que no pueden ser demostradas en él, pero cuyo carácter verdadero puede mostrarse por argumentos no formales, mediante la lógica de la metamatemática. Esta consecuencia, la de que existen limitaciones sobre lo que se puede conseguir mediante la axiomatización, contrasta nítidamente con la creencia de finales del siglo XIX de que la matemática concide exactamente con la colección de sus ramas axiomatizadas. Este resultado de Gödel asestó un golpe mortal a cualquier pretensión de axiomatizaciones globales. Esta incapacidad del método axiomático no es en sí misma una contradicción, pero resultó sorprendente porque los matemáticos esperaban que cualquier afir-

mación verdadera podría ser ciertamente demostrable dentro del marco de algún sistema axiomático. Desde luego, los argumentos anteriores no excluyen la posibilidad de nuevos métodos de demostración que permitan ir más allá de la metamatemática de Hilbert.

Hilbert no estaba convencido de que este golpe destruyera su programa. Sostenía que incluso aunque se necesitara usar conceptos exteriores a un sistema formal, aún podrían ser finitarios e intuitivamente concretos, y por tanto aceptables. Hilbert era profundamente optimista; tenía una confianza sin límites en el poder del entendimiento y del razonamiento humanos. En la conferencia que dio en el Congreso Internacional de 1928,[22] había dicho, «... no hay límites para el entendimiento matemático... en la matemática no hay ningún *Ignorabimus,* sino que siempre podremos contestar a las preguntas significativas... nuestra razón no posee ningún arte secreto, sino que procede de acuerdo con reglas bien definidas y estables que son la garantía de la objetividad absoluta de sus juicios.» Todo matemático, continúa, comparte la convicción de que cualquier problema matemático definido puede ser resuelto. Este optimismo le dio valor y fuerza pero le impidió entender que podrían presentarse problemas matemáticos indecidibles.

El programa formalista, tuviera éxito o no, resultaba inaceptable para los intuicionistas. En 1925 Brouwer atacaba enérgicamente a los formalistas.[23] Desde luego, decía, el tratamiento axiomático formalista evitará las contradicciones, pero de esta manera no se encontrará nada de valor matemático. Una teoría falsa no es menos falsa porque no conduzca a contradicción, lo mismo que un acto criminal es criminal esté o no condenado por un tribunal. Y añadía también sarcásticamente, «A la pregunta de dónde se encuentra el rigor matemático, las dos partes dan respuestas distintas. El intuicionista dice que en el intelecto humano; el formalista que en el papel.» Weyl atacó también el programa de Hilbert: «La matemática de Hilbert puede ser un bonito juego con fórmulas, más divertido aún que el ajedrez, pero qué relación tiene eso con el conocimiento, dado que se reconoce que sus fórmulas no tienen ningún significado material en virtud del cual pudieran expresar verdades intuitivas.» En defensa de la filosofía formalista hay que subrayar que únicamente con el

---

[22] Atti del Congresso Internazionale Dei Matematici, I, 135-141 = *Grundlagen der Geometrie*, 7.ª ed., 313-323.

[23] Jour. für Math., 154, 1925, 1.

fin de demostrar la consistencia, completitud y otras propiedades, se reduce la matemática a fórmulas carentes de significado. En cuanto a la matemática globalmente considerada, incluso la mayoría de los formalistas rechaza la idea de que se trate simplemente de un juego, sino que la consideran como una ciencia objetiva.

A su vez Hilbert acusó a Brouwer y a Weyl de intentar arrojar por la borda todo lo que no les convenía, promulgando de manera dictatorial un embargo,[24] y calificó el intuicionismo de traición a la ciencia. (Sin embargo, él mismo, en su metamatemática, se limitó a principios lógicos intuitivamente claros.)

8. Algunos desarrollos recientes

Ninguna de las soluciones propuestas a los problemas básicos de los fundamentos —la axiomatización de la teoría de conjuntos, el logicismo, el intuicionismo o el formalismo— lograron el objetivo de establecer un planteamiento de la matemática universalmente aceptable. Los desarrollos realizados desde la obra de Gödel de 1931 no han alterado esencialmente el cuadro. Sin embargo, ha habido unos pocos movimientos y resultados que vale la pena destacar. Algunos matemáticos han tratado de conseguir planteamientos de la matemática de compromiso, utilizando características de dos escuelas básicas. Otros, especialmente Gerhard Gentzen (1909-1945), miembro de la escuela de Hilbert, debilitaron las restricciones sobre los métodos de demostración permitidos en la metamatemática de Hilbert y así consiguió, por ejemplo, demostrar la consistencia de la aritmética y de algunas partes del análisis, utilizando inducción transfinita (inducción sobre los ordinales transfinitos numerables).[25]

Entre otros resultados importantes hay dos especialmente notables. En su *The Consistency of the Axiom of Choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory* (1940, ed. rev., 1951) demostró Gödel que si el sistema de axiomas de Zermelo-Fraenkel sin el axioma de elección es consistente, entonces el sistema obtenido añadiendo esta axioma también es consistente; es decir, el axioma de elección no puede ser refutado. Análogamente,

---

[24] Abh. Math. Seminar der Hamburger Univ., 1, 1922, 157-177 = *Ges. Abh.*, 3, 157-177.

[25] Math. Ann., 112, 1936, 493-565.

la hipótesis del continuo, de que no hay ningún número cardinal entre $\aleph$ y $2^{\aleph}$, es consistente con el sistema de Zermelo-Fraenkel (sin el axioma de elección). En 1963 Paul J. Cohen (1934), profesor de matemáticas de la universidad de Stanford, demostró [26] que estos dos últimos axiomas son independientes del sistema de Zermelo-Fraenkel; es decir, que no pueden ser demostrados dentro de dicho sistema. Además, incluso si se añade el axioma de elección al sistema de Zermelo-Fraenkel, la hipótesis del continuo sigue sin poder ser demostrada. Estos resultados implican que tenemos libertad para construir nuevos sistemas de matemáticas en los que se nieguen uno o los dos de estos controvertidos axiomas.

Todos los progresos realizados desde 1930 dejan abiertos dos importantes problemas: demostrar la consistencia del análisis clásico sin restricciones y de la teoría de conjuntos, y construir la matemática sobre una base intuicionista estricta o determinar los límites de este enfoque. El origen de las dificultades en estos dos problemas es el infinito utilizado tanto en el sentido de los conjuntos infinitos como en los procesos infinitos. Este concepto, que ya creó problemas a los griegos en conexión con los inconmensurables, y que eludieron mediante el método de exhausción, ha sido tema de discusión desde entonces, lo que hizo decir a Weyl que la matemática es en realidad la ciencia del infinito.

La cuestión acerca de la base lógica adecuada para la matemática y el nacimiento en particular del intuicionismo viene a sugerir que, en un sentido muy general, la matemática ha recorrido un círculo completo. Sus comienzos tuvieron una base intuitiva y empírica. El rigor se convirtió en una necesidad con los griegos y, aunque escasamente logrado hasta el siglo XIX, por un momento pareció alcanzado. Pero lo cierto es que todos los esfuerzos por perseguir el rigor hasta el final, han conducido a un *impasse* en el que ya no hay acuerdo acerca de qué es lo que realmente significa. La matemática sigue viva y con saludable vitalidad, pero sólo apoyándose en una base pragmática.

Algunos tienen la esperanza de que el aparente callejón sin salida actual tenga solución. El grupo de matemáticos franceses que escribieron con el seudónimo de Nicolás Bourbaki muestra su optimismo

---

[26] Proceedings of the National Academy of Sciences, 50, 1963, 1143-1148; 51, 1964, 105-110.

con las siguientes palabras:[27] «Hace veinticinco siglos que los matemáticos vienen practicando la costumbre de corregir sus errores, viendo así su ciencia enriquecida y no empobrecida; esto les da derecho a contemplar el futuro con serenidad.»

Esté justificado o no el optimismo, el estado actual de la matemática ha sido descrito adecuadamente por Weyl:[28] «El problema de los fundamentos últimos y del significado último de la matemática sigue abierto; no sabemos en qué dirección hallará su solución final, ni siquiera si cabe esperar en absoluto una respuesta final objetiva. El "matematizar" muy bien pudiera ser una actividad creativa del hombre, como el lenguaje o la música, de una originalidad primaria, cuyas decisiones históricas desafíen una racionalización objetiva completa.»

Bibliografía

Becker, Oskar: *Grundlagen der Mathematik in geschichtlicher Etwicklung*, Verlag Karl Alber, 1956, 317-401.

Beth, E. W.: *Mathematical Thought: An Introduction to the Philosophy of Mathematics*, Gordon and Breach, 1965.

Bochenski, I. M.: *A History of Formal Logic*, University of Notre Dame Press, 1962; Chelsea (reprint), 1970. Hay versión castellana, *Historia de la lógica formal*, Madrid, Gredos, 1985.

Boole, George: *An Investigation of the Laws of Thought* (1854), Dover (reprint), 1951. Hay versión castellana, *Una investigación sobre las leyes del pensamiento*, Madrid, Paraninfo, 1982.

— *The Mathematical Analysis of Logic* (1847), Basil Blackwell (reprint), 1948. Hay versión castellana, *El análisis matemático de la lógica*, Madrid, Cátedra, 1984.

— *Collected Logical Works*, Open Court, 1952.

Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.

Brouwer, L. E. J.: «Intuitionism and Formalism». *Amer. Math. Soc. Bull.*, 20, 1913/1914, 81-96. Traducción al inglés de la conferencia inaugural de Brouwer como profesor de matemáticas en Amsterdam.

Church, Alonzo: «The Richard Paradox». *Amer. Math. Monthly*, 41, 1934, 356-361.

---

[27] Journal of Symbolic Logic, 14, 1949, 2-8.

[28] Obituary Notices of Fellows of the Royal Soc., 4, 1944, 547-553 = *Ges. Abh.*, 4, 121-129, parte de la p. 126.

Cohen, Paul J., y Reuben Hersh: «Non-Cantorian Set Theory». *Scientific American*, Dec. 1967, 104-116.

Couturat, L.: *La Logigue de Leibniz d'après des documents inédits*, Alcan, 1901.

De Morgan, Augustus: *On the Syllogism and Other Logical Writings*, Yale University Press, 1966. Una colección de sus artículos, editada por Peter Heath.

Dresden, Arnold: «Brouwer's Contribution to the Foundations of Mathematics». *Amer. Math. Soc. Bull.*, 30, 1924, 31-40.

Enriques, Federigo: *The Historic Development of Logic*, Henry Holt, 1929.

Fraenkel, A. A.: «The Recent Controversies About the Foundations of Mathematics». *Scripta Mathematica*, 13, 1947, 17-36.

Fraenkel, A. A., y Bar-Hillel, Y.: *Foundations of Set Theory*, North-Holland, 1958.

Frege, Gottlob: *The Foundations of Arithmetic*, Blackwell, 1953, en inglés y en alemán; también la traducción inglesa sola, Harper and Bros., 1960. Hay versión castellana, *Fundamentos de la aritmética*, Barcelona, Laia, 1972.

— *The Basic Laws of Arithmetic*, University of California Press, 1965.

Gerhardt, C. I. (ed.): *Die philosophischen Schriften von G. W. Leibniz*, 1875-1880, vol. 7.

Gödel, Kurt: *On Formally Undecidable Propositions of* Principia Mathematica *and Related Systems*, Basic Books, 1965. Hay versión castellana en K. Gödel, *Obras completas*, Madrid, Alianza, 1981, pp. 55-89.

— «What Is Cantor's Continuum Problem?», *Amer. Math. Monthly*, 54, 1947, 515-525. Hay versión castellana, *ibid.*, pp. 340-362.

— *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press, 1940; rev. ed., 1951.

Kneale, William, y Martha: *The Development of Logic*, Oxford University Press, 1962. Hay versión castellana, W. y H. Kneale, El desarrollo de la lógica, Madrid, Tecnos, 1972.

Kneebone, G. T.: *Mathematical Logic and the Foundations of Mathematics*, D. Van Nostrand, 1963. *Véase especialmente el apéndice sobre los desarrollos desde 1939.*

Leibniz, G. W.: *Logical Papers*, traducido y editado por G. A. R. Parkinson, Oxford University Press, 1966.

Lewis, C. I.: *A Survey of Symbolic Logic*, Dover (reprint), 1960, 1-117.

Meschkowski, Herbert: *Probleme des Unendlichen, Werk und Leben Georg Cantors*, F. Vieweg und Sohn, 1967.

Mostowski, Andrzej: *Thirty Years of Foundational Studies*, Barnes and Noble, 1966.

Nagel, E., y Newman, J. R.: *Gödel's Proof*, New York University Press, 1958. Hay versión castellana, *El teorema de Gödel*, Madrid, Tecnos, 1979.

Poincaré, Henri: *The Foundations of Science*, Science Press, 1946, 448-485. Esta es una reimpresión en un volumen de *Science and Hypothesis*, *The Value of Science*, and *Science and Method. Hay versiones castellanas, La ciencia y la hipótesis, El valor de la ciencia, Ciencia y método*, Madrid, Espasa Calpe.

Rosser, J. Barkley: «An Informal Exposition of Proofs of Gödel's Theorems and Church's Theorem». *Journal of Symbolic Logic*, 4, 1939, 53-60.

Russell, Bertrand: *The Principles of Mathematics*, George Allen and Unwin, 1903; 2.ª ed., 1937. Hay versión castellana, *Los principios de la matemática*, Madrid, Espasa Calpe, 4.ª ed., 1983.

Scholz, Heinrich: *Concise History of Logic*, Philosophical Library, 1961.

Styazhkin, N. I.: *History of Mathematical Logic from Leibniz to Peano*, Massachusetts Institute of Technology Press, 1969.

Van Heijenoort, Jean: *From Frege to Gödel*, Harvard University Press, 1967. Traducciones de artículos clave sobre lógica y fundamentos de la matemática.

Weyl, Hermann: «Mathematics and Logic». *Amer. Math. Monthly*, 53, 1946, 2-13 = *Ges. Abh.*, 4, 268-279.

— *Philosophy of Mathematics and Natural Science*, Princeton University Press, 1949.

Wilder, R. L.: «The Role of the Axiomatic Method». *Amer. Math. Monthly*, 74, 1967, 115-127.

INDICE ONOMASTICO

INDICE DE MATERIAS